aruba®

NETWORKS

Implementing ArubaOS-CX Switching Volume 2

TRAINING MANUAL

EDU-ICX-ILT-v20.211

Implememting ArubaOS-CX Switching Student Guide

Open Source Code
Certain Aruba products include Open Source software code developed by third parties, including software code subject to the GNU General Public License ("GPL"), GNU Lesser General Public License ("LGPL"), or other Open Source Licenses. The Open Source code used can be found at this site:
http://www.arubanetworks.com/open_source

Legal Notice
The use of Aruba Networks, Inc. switching platforms and software, by all individuals or corporations, to terminate other vendors' VPN client devices constitutes complete acceptance of liability by that individual or corporation for this action and indemnifies, in full, Aruba Networks, Inc. from any and all legal actions that might be taken against it with respect to infringement of copyright on behalf of those vendors.

Warranty
This hardware product is protected by the standard Aruba warranty of one year parts/labor. For more information, refer to the ARUBACARE SERVICE AND SUPPORT TERMS AND CONDITIONS.
Altering this device (such as painting it) voids the warranty.

SKU: EDU-ICX-ILT-
v20.21 April 2020

Implementing ArubaOS-CX Switching Volume 2

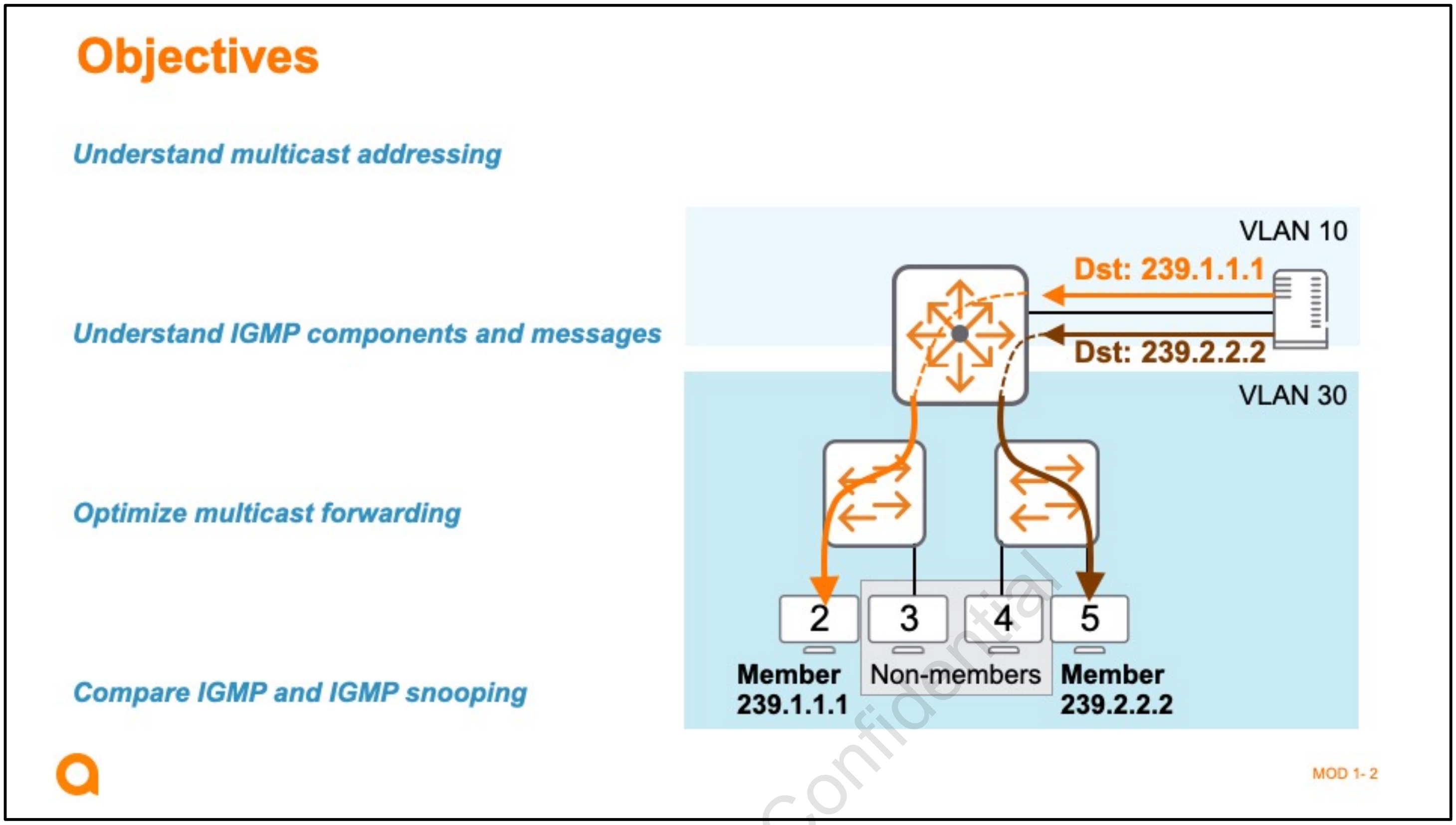
Table of Contents

Rev # 20.21

IGMP

Implementing ArubaOS-CX Switching

Welcome to Module 8 – IGMP.



Objectives

Understand multicast addressing

Understand IGMP components and messages

Optimize multicast forwarding

Compare IGMP and IGMP snooping

MOD 1- 2

After completing this module, you will be able to:

Understand multicast addressing

Understand IGMP components and messages

Optimize multicast forwarding

Compare IGMP and IGMP snooping

Overview

Multicast Intro

IGMP

Lab Activity

MOD 1- 3

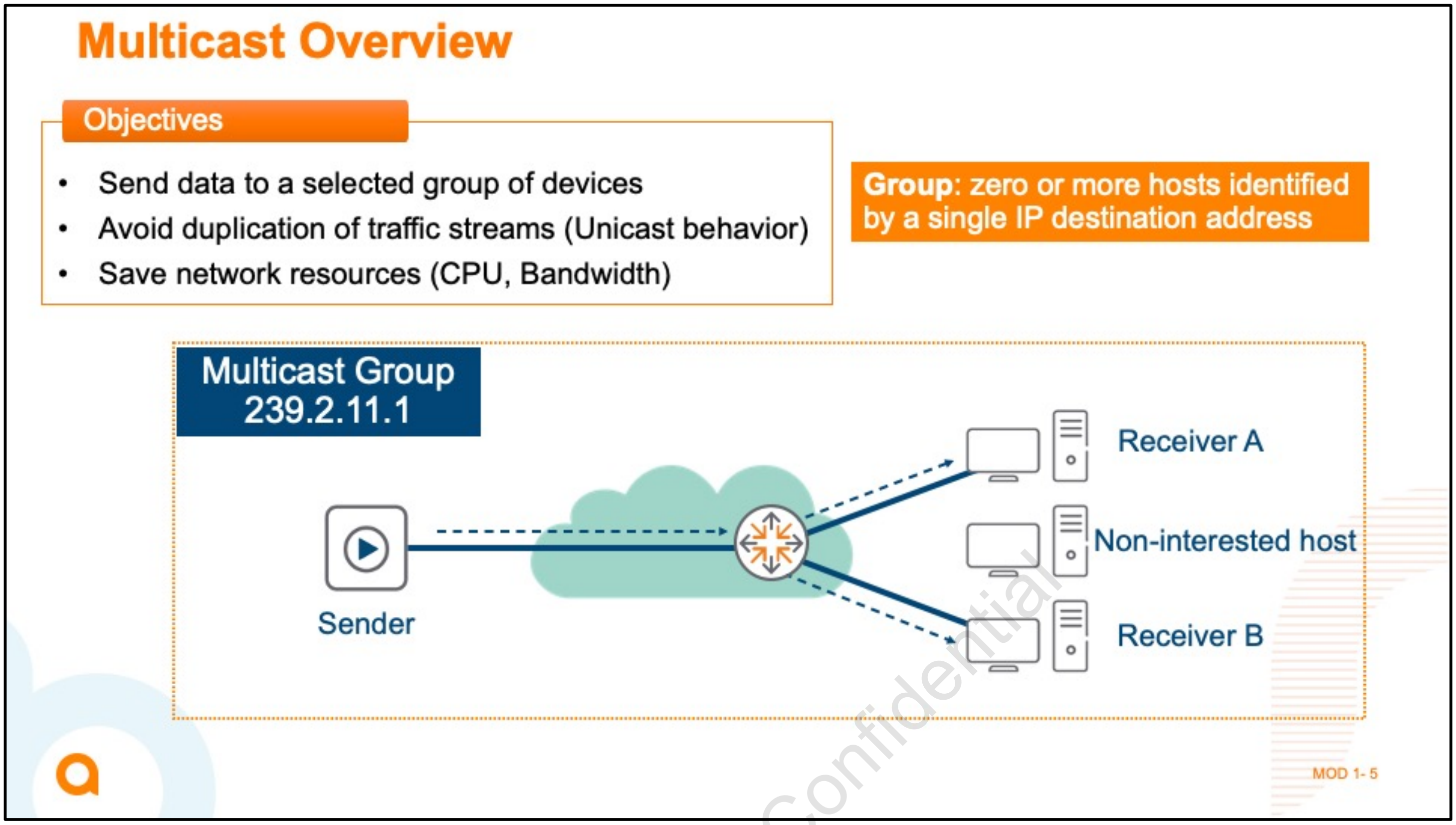
Enterprise networks require a variety of multicast services. In this module, you learn about multicast traffic and how to use Internet Group Management Protocol (IGMP) to control it.

We begin with an introduction to multicasting, then move into a discussion of IGMP before turning you loose on a lab activity.

Multicast Intro

aruba
a Hewlett Packard Enterprise company

We begin with an introduction to multicasting.



Multicast Overview

Objectives

- Send data to a selected group of devices
- Avoid duplication of traffic streams (Unicast behavior)
- Save network resources (CPU, Bandwidth)

Group: zero or more hosts identified by a single IP destination address

Multicast Group 239.2.11.1

MOD 1- 5

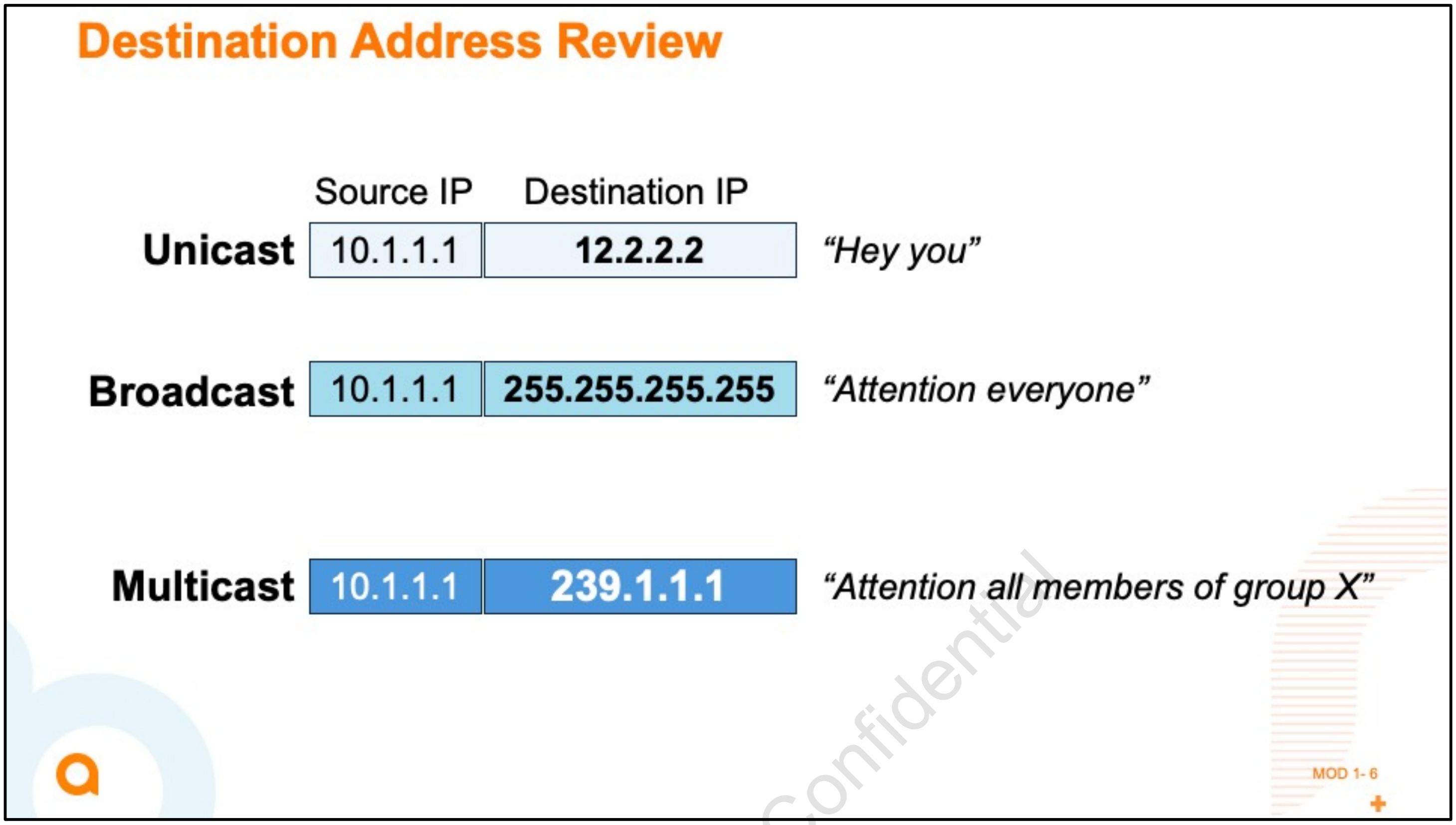
The purpose of multicast is to send information to a select group of devices and to avoid duplication of traffic streams. This saves precious network resources such as bandwidth and CPU utilization.

A sender device generates information and advertises it to a group. Receivers join that group and listen to the information. A multicast group is defined as a set of zero or more host identified by a single IP destination address. The figure shows a multicast group represented by IP address 239.2.11.1.

IP multicast messages are transmitted from one-to-many, from many-to-many, and/or from many-to-one over an IP infrastructure that may span Layer 3 boundaries. The destination nodes (receivers) send join and leave messages that create an on-demand community of devices interested in receiving the multicast stream. Multicast optimally uses network resources by requiring the source to send a single packet stream, even though large numbers of receivers might need to receive the messages. The replication of messages takes place at the network node(s) or Layer 3 device(s) and diverges the stream to the receivers.

The multicast source sends packets with its own IP address as the source, and some multicast address as the destination. Multiple receivers can listen for the same address, and so switches flood the packet out all ports in the same VLAN, except the packet ingress port. Any source that is listening for that address processes the packet. Because multiple IPv4 nodes listen for the same multicast address, this address is often referred to as a group address.

Note: The term stream is commonly used in multicast to denote traffic associated with a particular multicast destination address.



Destination Address Review

	Source IP	Destination IP	
Unicast	10.1.1.1	**12.2.2.2**	*"Hey you"*
Broadcast	10.1.1.1	**255.255.255.255**	*"Attention everyone"*
Multicast	10.1.1.1	**239.1.1.1**	*"Attention all members of group X"*

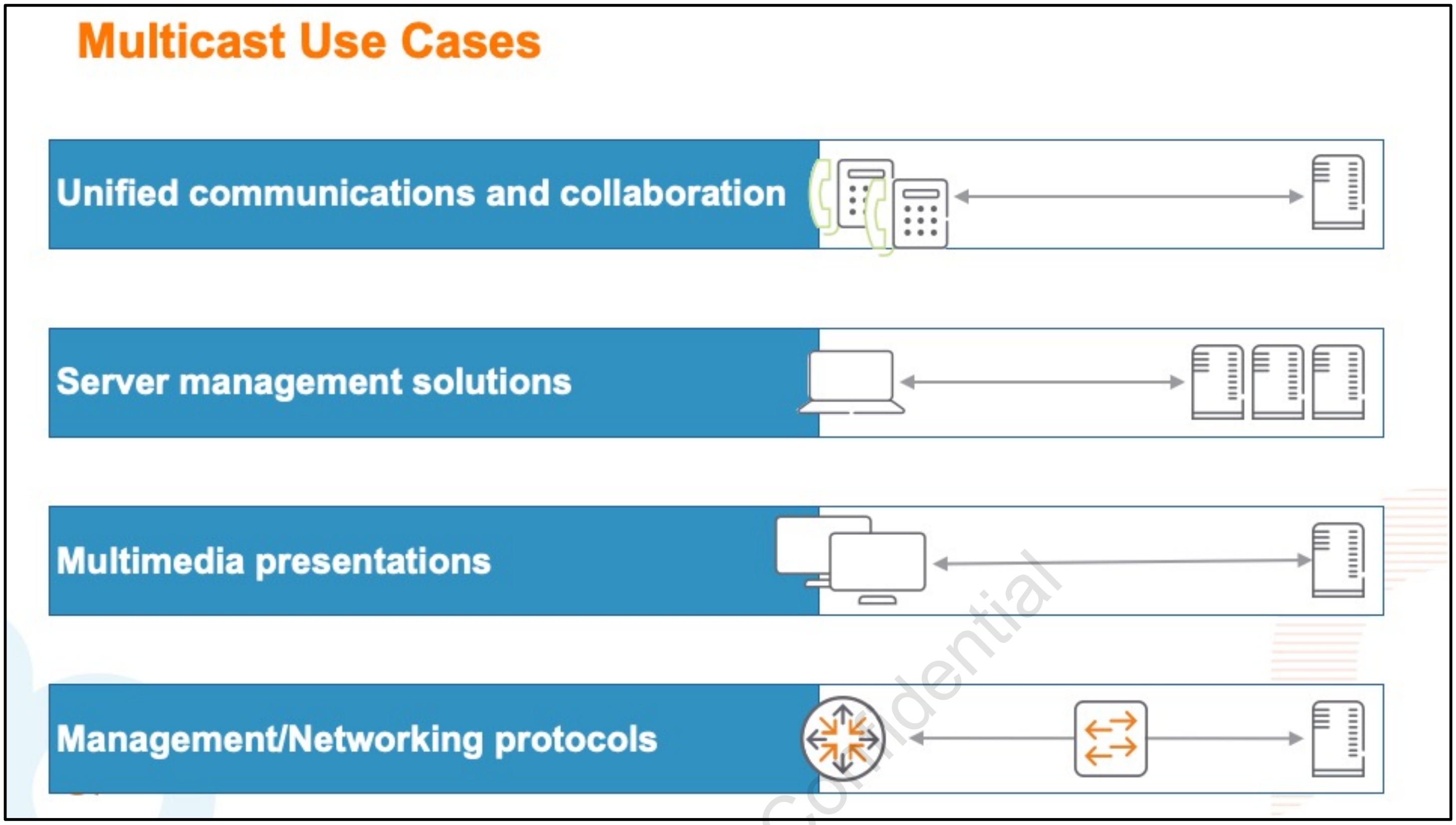
MOD 1- 6

As a review, there are three general types of destination addresses:

| Unicast: traffic is forwarded to just one destination

| Broadcast: traffic is flooded throughout the VLAN/subnet (everyone will see it)

| Multicast: traffic is destined to 0 or more (or all) destinations in a VLAN/subnet. It is flooded, by default, throughout the VLAN/subnet



Multicast Use Cases

Unified communications and collaboration

Server management solutions

Multimedia presentations

Management/Networking protocols

Unified communications and collaboration(UCC)

UCC applications include the following:

- IP voice conference calling
- Music or audio streaming
- Video and desktop conferencing and collaboration
- Telepresence
- Video streaming applications
- Virtual classroom applications
- Stock ticker tape streaming applications
- Network management protocols like the Network Time Protocol (NTP) and the Aruba Discovery Protocol (ADP), used by Aruba campus APs to find and connect to an Aruba Mobility Controller

Server management solutions

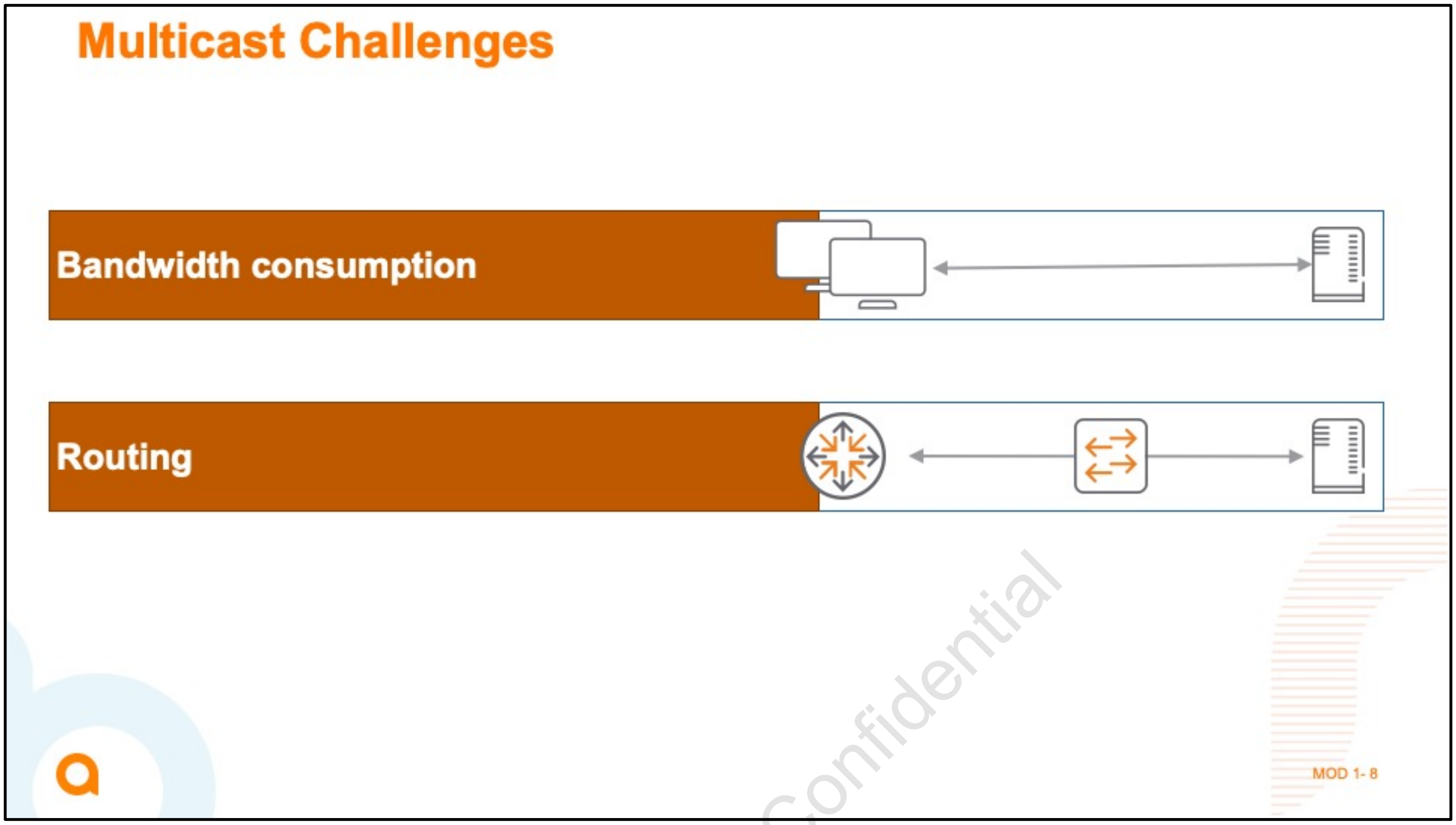
IT administrators might also assign servers or endpoints to multicast groups so that they can deliver patches, software updates, and other controls to a targeted group. Multicasting can also be used to take server or workstations snapshots – images of each syst em's storage drives.

Multimedia presentations

This is often the first thing folks think about when discussing multicasting. Videos and other multimedia presentations can be large files, and consume significant bandwidth. If 1000 people needed to watch the CEO's quarterly presentation to staff, you do not want 1000 individual unicast data streams. It is far better to send one single stream as a multicast, such that all staff members receive it.

Management/networking protocols

Many network-related protocols use multicasting for efficiency. This includes the Network Time Protocol (NTP), Aruba Discovery Protocol (ADP), OSPF, and many more.



Multicast Challenges

Bandwidth consumption

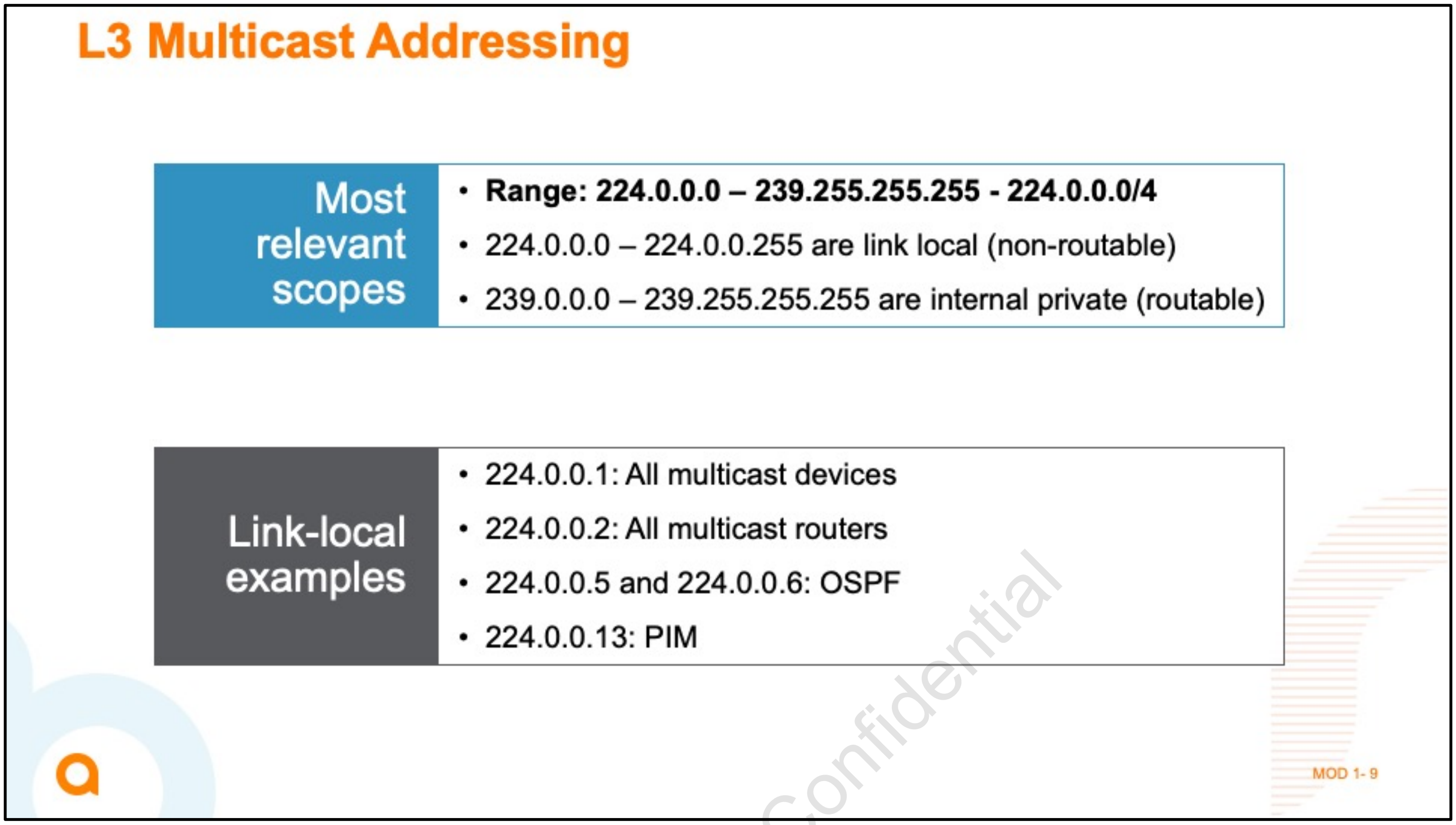
Routing

MOD 1- 8

Multicast traffic for sophisticated applications, particularly multimedia applications, can consume a great deal of bandwidth, particularly when it is copied for distribution to multiple locations. In fact, even copying multicasts can consume processing power on network infrastructure devices. To maintain a high-performing network environment, administrators must implement multicast control technologies properly to minimize unnecessary flooding and duplication. Not only do administrators need to preserve bandwidth, they also need to protect endpoint functionality. An endpoint NIC must process the multicasts that it receives, even if it is not listening for that multicast, which can consume CPU cycles.

Users and devices that require specific multicast traffic often connect in a variety of locations. Network administrators might need to route the traffic across multiple VLANs and even campus LAN sites.

You will learn about meeting this first challenge in this module. Then, in the next module, you will build on that knowledge and learn about multicast routing.



L3 Multicast Addressing

Most relevant scopes	• **Range: 224.0.0.0 – 239.255.255.255 - 224.0.0.0/4** • 224.0.0.0 – 224.0.0.255 are link local (non-routable) • 239.0.0.0 – 239.255.255.255 are internal private (routable)
Link-local examples	• 224.0.0.1: All multicast devices • 224.0.0.2: All multicast routers • 224.0.0.5 and 224.0.0.6: OSPF • 224.0.0.13: PIM

MOD 1- 9

Multicast addresses fall within the 224.0.0.0/4 range - 224.0.0.0 to 239.255.255.255. As shown in the figure, the lower portion of the range is reserve for link local addresses – they are non-routable addresses that are used to communicate among devices on the same subnet. The remaining addresses are internal private addresses – routable multicast addresses for use across the enterprise.

The Internet Assigned Numbers Authority (IANA) has assigned specific multicast addresses to some protocols. For example, to reach all multicast host devices, use 224.0.0.1. To reach all routers, use 224.0.0.2. Open Shortest Path First (OSPF) routers communicate with each other on multicast address 224.0.0.5 and 224.0.0.6. 224.0.0.13 is used for the Protocol Independent Multicast (PIM) routing protocol.

Depending on the type of application, the multicast source and multicast receivers might be statically configured to stream to, and listen for, a specific multicast address. Or they might negotiate the address dynamically. Determining the multicast address to use is the application’s responsibility; the network infrastructure must simply ensure that the traffic reaches its destination.

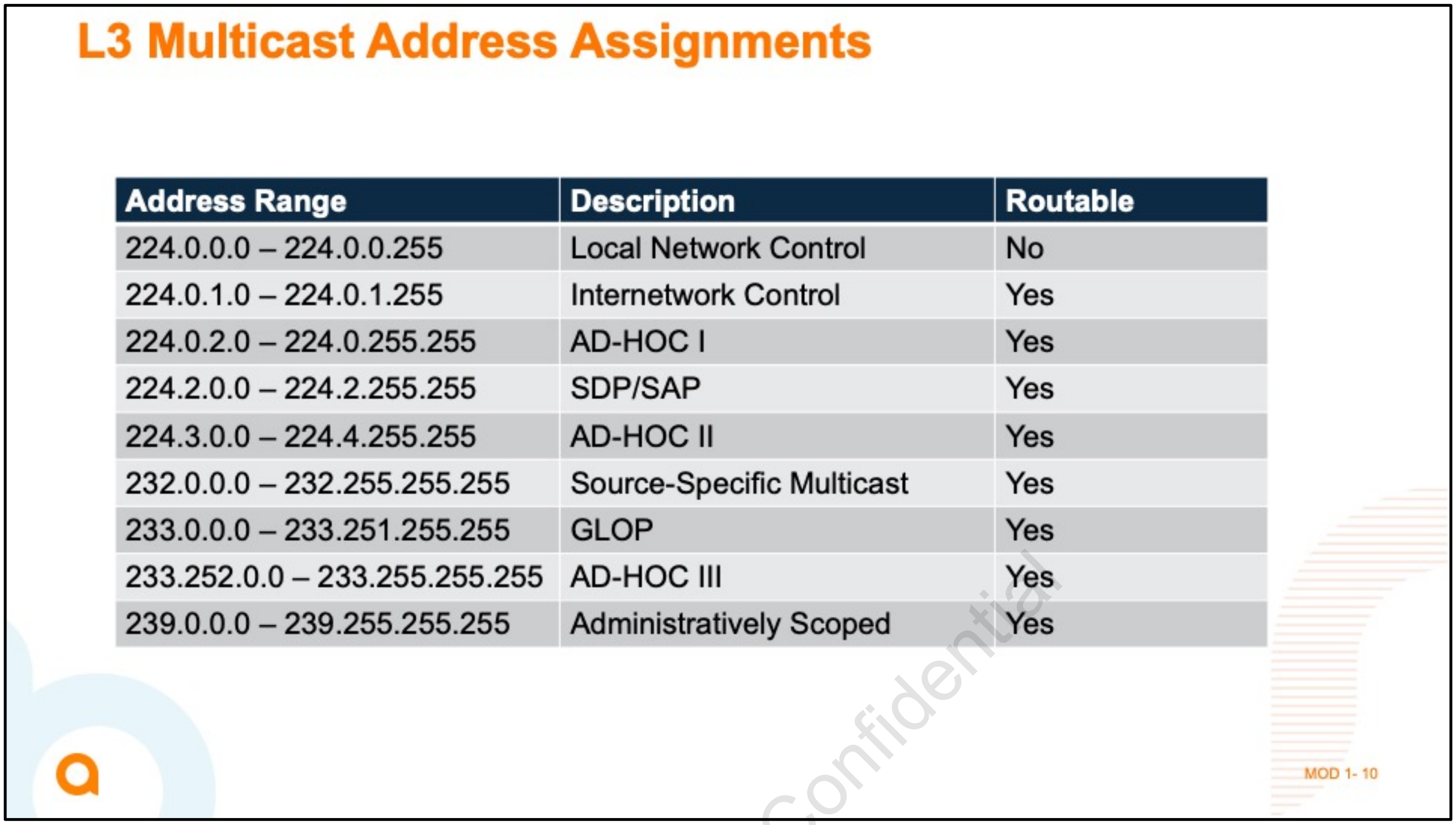
As shown in the table below, IANA has assigned other addresses to specific organizations.

Some applications, such as applications that use Session Announcement Protocol (SAP), can dynamically select multicast addresses from a certain range. Finally, individual organizations can define administratively scoped multicast addresses in the 239.0.0.0/8 range as they want; these multicasts should not leave the administrative domain.

This table summarizes the Layer-3 multicast addresses:

IP multicast address range	Description	Routable
224.0.0.0 to 224.0.0.255	Local subnetwork	No
224.0.1.0 to 224.0.1.255	Internetwork control	Yes
224.0.2.0 to 224.0.255.255	AD-HOC block 1	Yes
224.3.0.0 to 224.4.255.255	AD-HOC block 2	Yes
232.0.0.0 to 232.255.255.255	Source-specific multicast	Yes
233.0.0.0 to 233.251.255.255	GLOP addressing	Yes
233.252.0.0 to 233.255.255.255	AD-HOC block 3	Yes
234.0.0.0 to 234.255.255.255	Unicast-prefix-based	Yes
239.0.0.0 to 239.255.255.255	Administratively scoped	Yes

You can refer to IANA for the latest assignments: https://www.iana.org/assignments/multicast-addresses/multicast-addresses.xhtml.



L3 Multicast Address Assignments

Address Range	Description	Routable
224.0.0.0 – 224.0.0.255	Local Network Control	No
224.0.1.0 – 224.0.1.255	Internetwork Control	Yes
224.0.2.0 – 224.0.255.255	AD-HOC I	Yes
224.2.0.0 – 224.2.255.255	SDP/SAP	Yes
224.3.0.0 – 224.4.255.255	AD-HOC II	Yes
232.0.0.0 – 232.255.255.255	Source-Specific Multicast	Yes
233.0.0.0 – 233.251.255.255	GLOP	Yes
233.252.0.0 – 233.255.255.255	AD-HOC III	Yes
239.0.0.0 – 239.255.255.255	Administratively Scoped	Yes

MOD 1- 10

As shown in the figure, IANA has assigned other addresses to specific organizations and uses. Some applications, such as applications that use Session Announcement Protocol (SAP), can dynamically select multicast addresses from a certain range. Finally, individual organizations can define administratively scoped multicast addresses in the 239.0.0.0/8 range as they want; these multicasts should not leave the administrative domain.

The group address simply identifies a group of nodes interested in a flow. The group address combined with the source IP address identifies the multicast flow.

The table identifies the Multicast Address Assignment:

Local Network Control block (224.0.0/24): The local control block is used for specific protocol control traffic. Router interfaces listen to but do not forward local control multicasts; for example, OSPF “all routers” (224.0.0.5). Assignments in this block are publicly controlled by IANA. You can find a complete list of Local Network Control address assignments at the IANA website (www.iana.org).

Internetwork Control block (224.0.1/24): The Internetwork Control block is for protocol control traffic that router interfaces may forward through the Autonomous System Number (ASN) or through the Internet. Examples include 224.0.1.1 Network Time Protocol (NTP), defined in RFC 4330, and 224.0.1.68 mdhcpdiscover, defined in RFC 2730. Internetwork Control group assignments are also publicly controlled by IANA.

AD-HOC blocks (I: 224.0.2.0–224.0.255.255, II: 224.3.0.0–224.4.255.255, and III:233.252.0.0–233.255.255.255): Traditionally assigned to applications that do not fit in either the Local or Internetwork Control blocks. Router interfaces may forward AD-HOC packets globally. Most applications using AD-HOC blocks require few group addresses (such as, for example, less than a /24 space). IANA controls any public AD-HOC Block assignments.

SDP/SAP block (224.2.0.0/16): The Session Description Protocol/Session Announcement Protocol (SDP/SAP) block is assigned to applications that receive addresses through the SAP as described in RFC 2974.

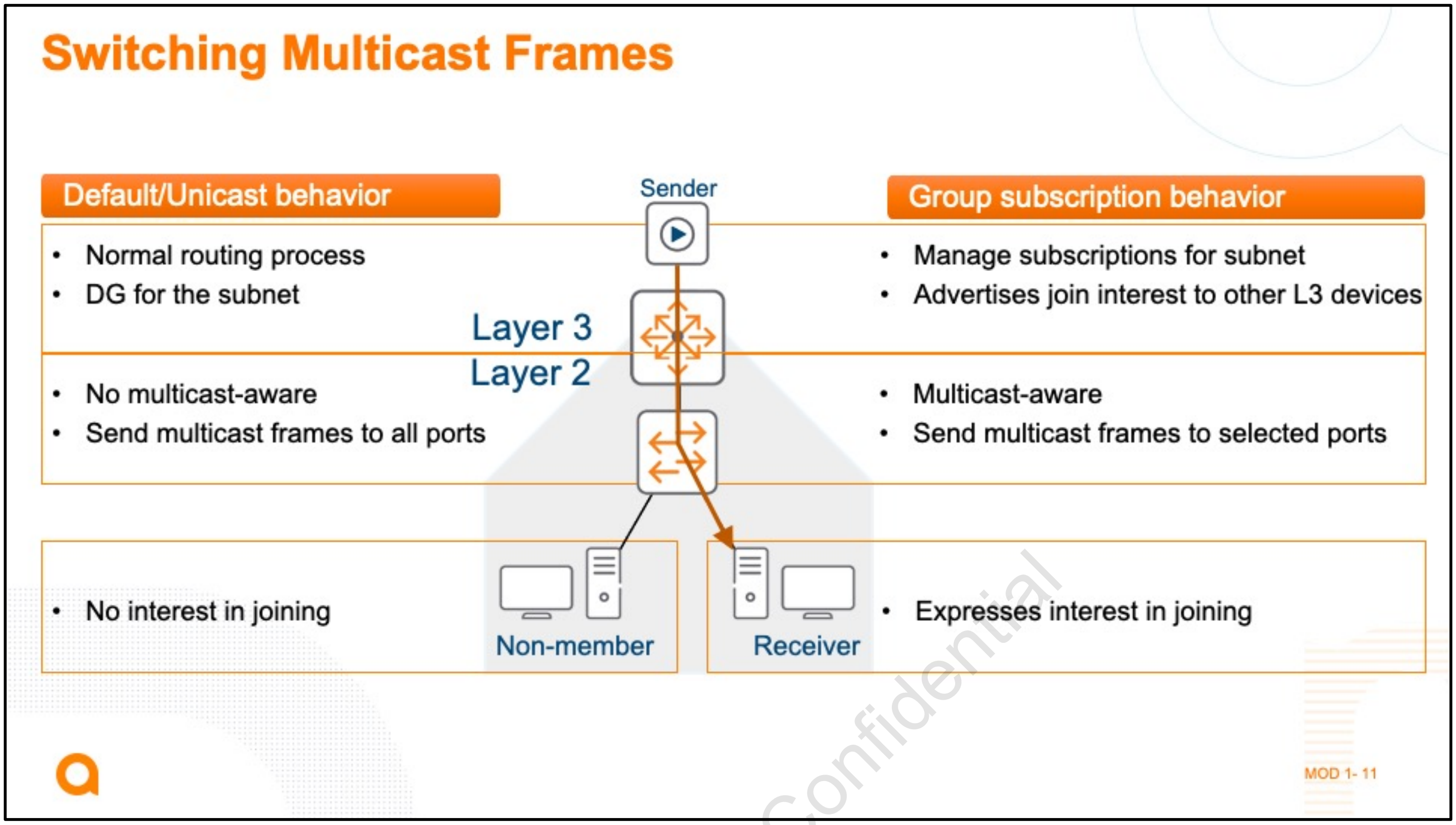
Source-Specific Multicast block (232.0.0.0/8): SSM addressing is defined by RFC 4607. SSM is a group model of IP Multicast in which multicast traffic is forwarded to receivers from only those multicast sources for which the receivers have explicitly expressed interest. SSM is mostly used in one-to-many applications. No official assignment from IANA is required to use the SSM block because the application is local to the host; however, according to IANA policy, the block is explicitly reserved for SSM applications and must not be used for any other purpose.

GLOP block (233.0.0.0/8): These addresses are statically assigned with a global scope. Each GLOP static assignment corresponds to a domain with a public 16-bit autonomous system number (ASN), which is issued by IANA. The ASN is inserted in dotted-decimal into the middle two octets of the multicast group address (X.Y). An example GLOP assignment for an ASN of X.Y would be 233.X.Y.0/24. Domains using an assigned 32-bit ASN should apply for group assignments from the AD-HOC III Block. Another alternative is to use IPv6 multicast group addressing. Because the ASN is public, IANA does not need to assign the actual GLOP groups. The GLOP Block is intended for use by public content, network, and Internet service providers. IANA considers GLOP addressing to be experimental, and 233.252–255.0.0 is reserved.

Administratively Scoped block (239.0.0.0/8): Administratively Scoped addresses are intended for local use within a private domain as described by RFC 2365. These group addresses serve a similar function as RFC 1918 private IP address block (such as, for example, 10.0.0.0/8 or 172.16-31.0.0/16 blocks). Network architects can create an address schema using this block that best suits the needs of the private domain and can further split scoping into specific geographies, applications, or networks.

Note: “GLOP”, surprisingly, is not an acronym. It does not stand for anything. The authors of the RFC needed a better way to refer to “that addressing method where you put your AS in the middle two octets”. David Meyers started calling it “GLOP” and the name stuck.

You can refer to IANA for the latest assignments: https://www.iana.org/assignments/multicast-addresses/multicast-addresses.xhtml.



Switching Multicast Frames

Default/Unicast behavior

- Normal routing process
- DG for the subnet

- No multicast-aware
- Send multicast frames to all ports

- No interest in joining

Group subscription behavior

- Manage subscriptions for subnet
- Advertises join interest to other L3 devices

- Multicast-aware
- Send multicast frames to selected ports

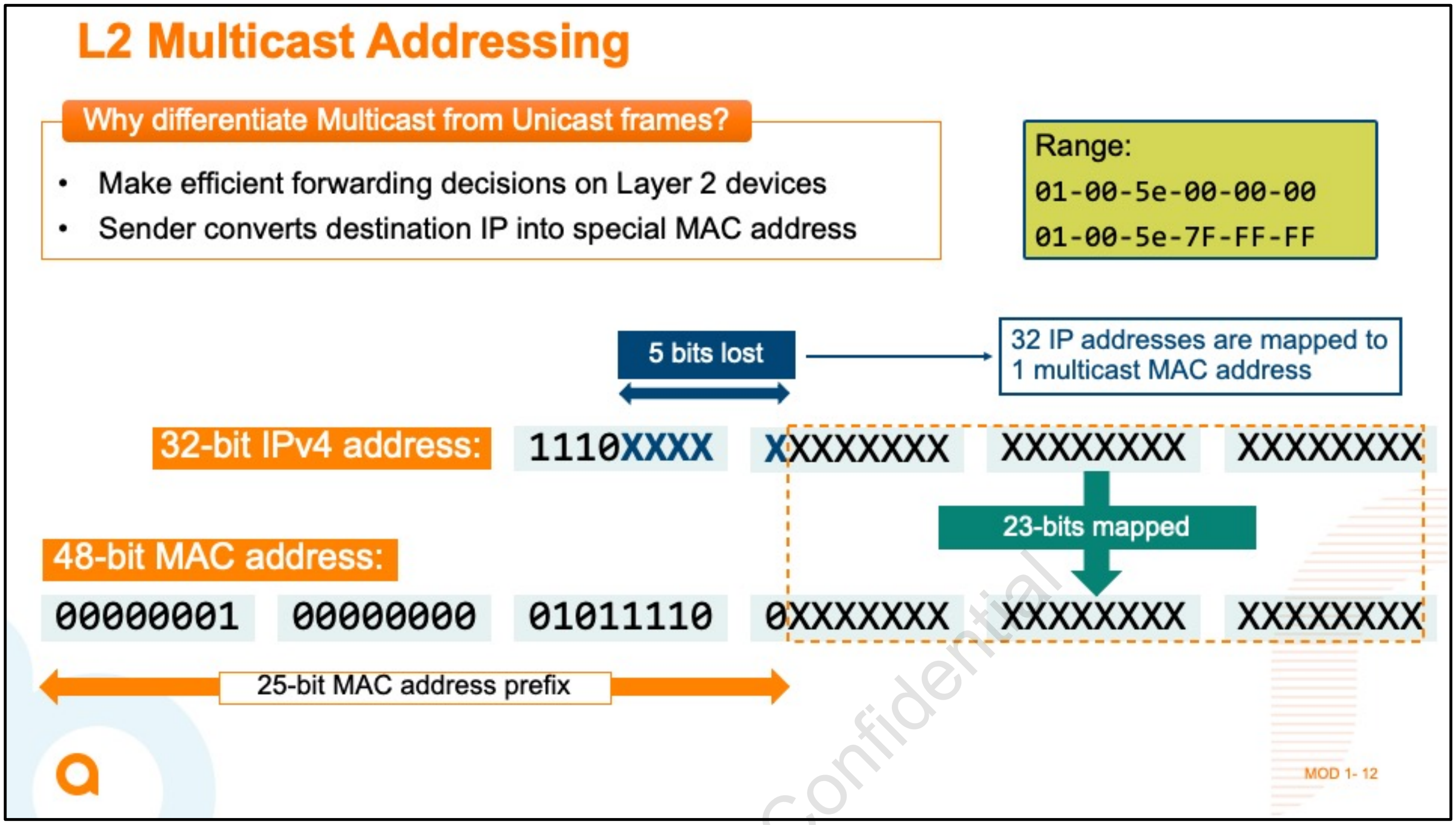
- Expresses interest in joining

MOD 1- 11

The major difference between a unicast and multicast network is that, in a unicast network all traffic is fully permitted to the host, assuming no filter is applied. In a multicast network, the host must advertise its intention to join the network.

Without multicast-aware Layer 2 protocols, all hosts on a given Layer 2 segment will receive multicast packets for any groups joined by a host on that segment. This is not efficient, since non-interested endpoints must process each frame, only to drop it. When Layer 2 devices are multicast-aware they properly send multicast traffic to only the ports where we find interested hosts.

The gateway is the network demarcation between Layers 2 and 3 and is the most appropriate device to manage host group membership for the larger network. This device receives these management messages and adds host segment interfaces to the local multicast table. A gateway communicates group interest for a multicast group to other Layer 3 devices using multicast routing (PIM).



L2 Multicast Addressing

Why differentiate Multicast from Unicast frames?

- Make efficient forwarding decisions on Layer 2 devices
- Sender converts destination IP into special MAC address

Range:
01-00-5e-00-00-00
01-00-5e-7F-FF-FF

5 bits lost → 32 IP addresses are mapped to 1 multicast MAC address

32-bit IPv4 address: 1110XXXX XXXXXXXX XXXXXXXX XXXXXXXX

23-bits mapped

48-bit MAC address: 00000001 00000000 01011110 0XXXXXXX XXXXXXXX XXXXXXXX

25-bit MAC address prefix

To optimize network resources, an Ethernet switch must understand multicast addressing. This is where the magic happens. The sending device must convert the destination IP multicast address into a special MAC address as follows:

The high-order 25 bits is the official reserved multicast MAC address range from 0100.5E00.0000 to 0100.5E7F.FFFF (RFC-1112). These bits are part of the organizational unit identifiers (OUI).

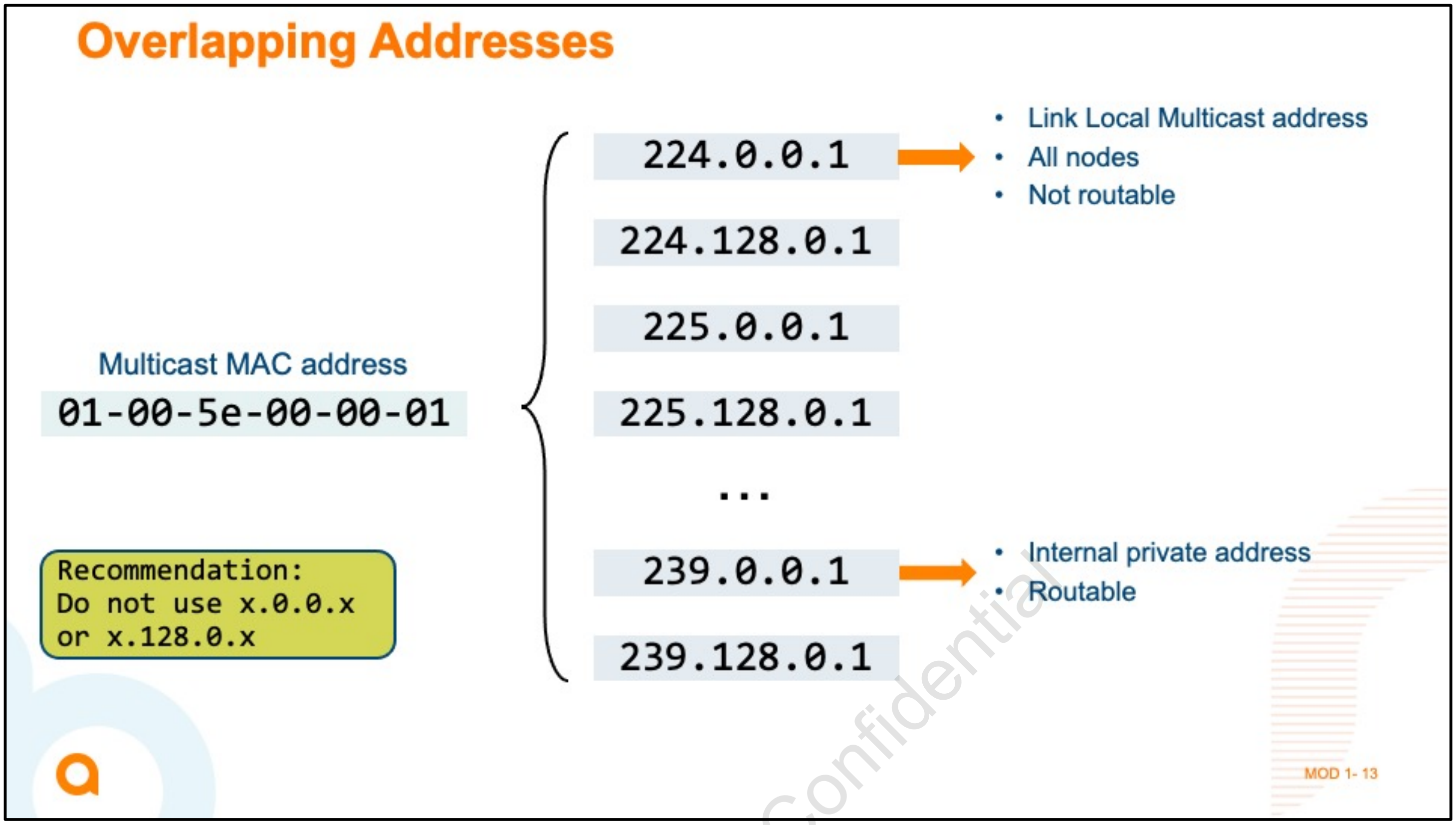
The lower-order 23 bits of the destination IP multicast address are mapped to the lower-order 23 bits of the MAC address.

A switch can use this calculated multicast MAC address to distinguish a frame as a multicast and make efficient forwarding decisions

We miss 5 bits of mapping information: 25 = 32. This means we will map 32 multicast IP addresses to 1 multicast MAC address.

The IEEE has allocated the address block 01-80-C2-00-00-00 to 01-80-C2-FF-FF-FF for group addresses for use by standard protocols implementing multicast at Layer-2. Of these, the MAC group addresses in the range of 01-80-C2-00-00-00 to 01-80-C2-00-00-0F are not relayed by switches conforming to 802.1D; i.e., they are not flooded. Examples of these would include LACP, STP BPDUs, LLDP, CDP, and others.

Important: With a broadcast, all devices in a subnet receive it. Their respective NICs then have the local device’s CPU process it. With a multicast, the NIC can make an intelligent decision as to whether the CPU should process the frame or if the NIC should drop the frame based on if the multicast MAC address is in use by a multicast-based application on the device. Therefore, the NIC can conserve CPU cycles based on what multicast addresses need to be seen by a running multicast application on the device, making multicast a much better solution to deliver information to a broad range of devices than broadcasts.



Overlapping Addresses

Multicast MAC address

01-00-5e-00-00-01

224.0.0.1:

- Link Local Multicast address
- All nodes
- Not routable

239.0.0.1:

- Internal private address
- Routable

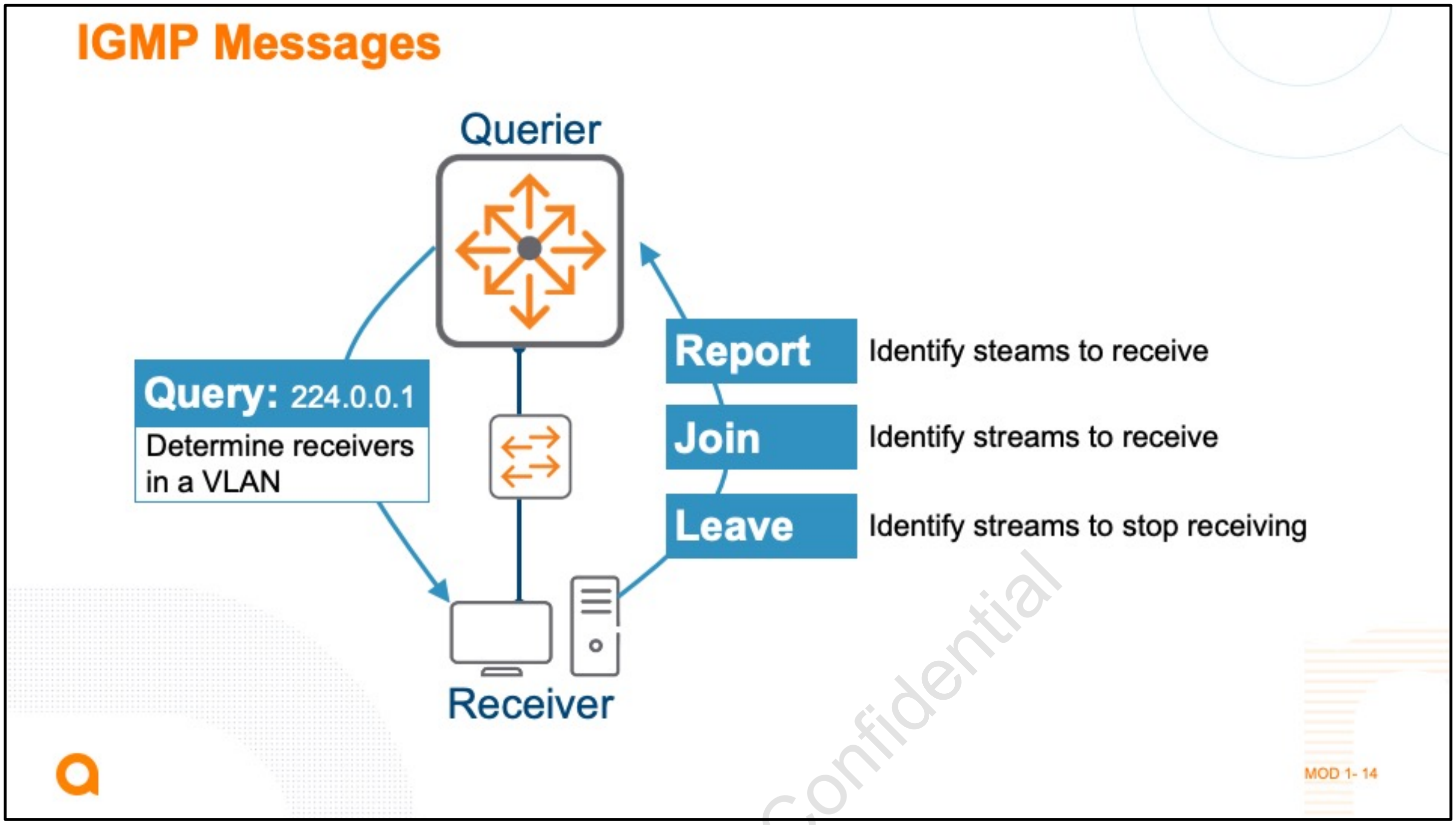
Recommendation:
Do not use x.0.0.x
or x.128.0.x

The multicast IP addresses above all map to the same multicast MAC address (01-00-5E-00-00-01).

This can cause some network problems. For example, a host that listens to the 239.0.0.1 multicast IP address will configure its network card to listen to MAC address 01-00-5E-00-00-01. At the same time, any device could generate traffic to 224.0.0.1 (all-nodes in the link) which is mapped to the same MAC address hosts will have to look at the IP address (Process Layer 3 header) of the received frame to see if it's for 239.0.0.1 and discard frames that are meant for 224.0.0.1. This process is not efficient.

As a recommendation do not use x.0.0.x or x.128.0.x, since these addresses will overlay with the Link-Local Multicast address scope.

Note. Some network devices do IGMP filtering based on MAC address. This implies that you should be careful when picking your multicast groups that you do not overlay the mac-addresses as to avoid odd interactions at layer-2.



IGMP Messages

Querier

Query: 224.0.0.1
Determine receivers in a VLAN

Report Identify steams to receive

Join Identify streams to receive

Leave Identify streams to stop receiving

Receiver

MOD 1- 14

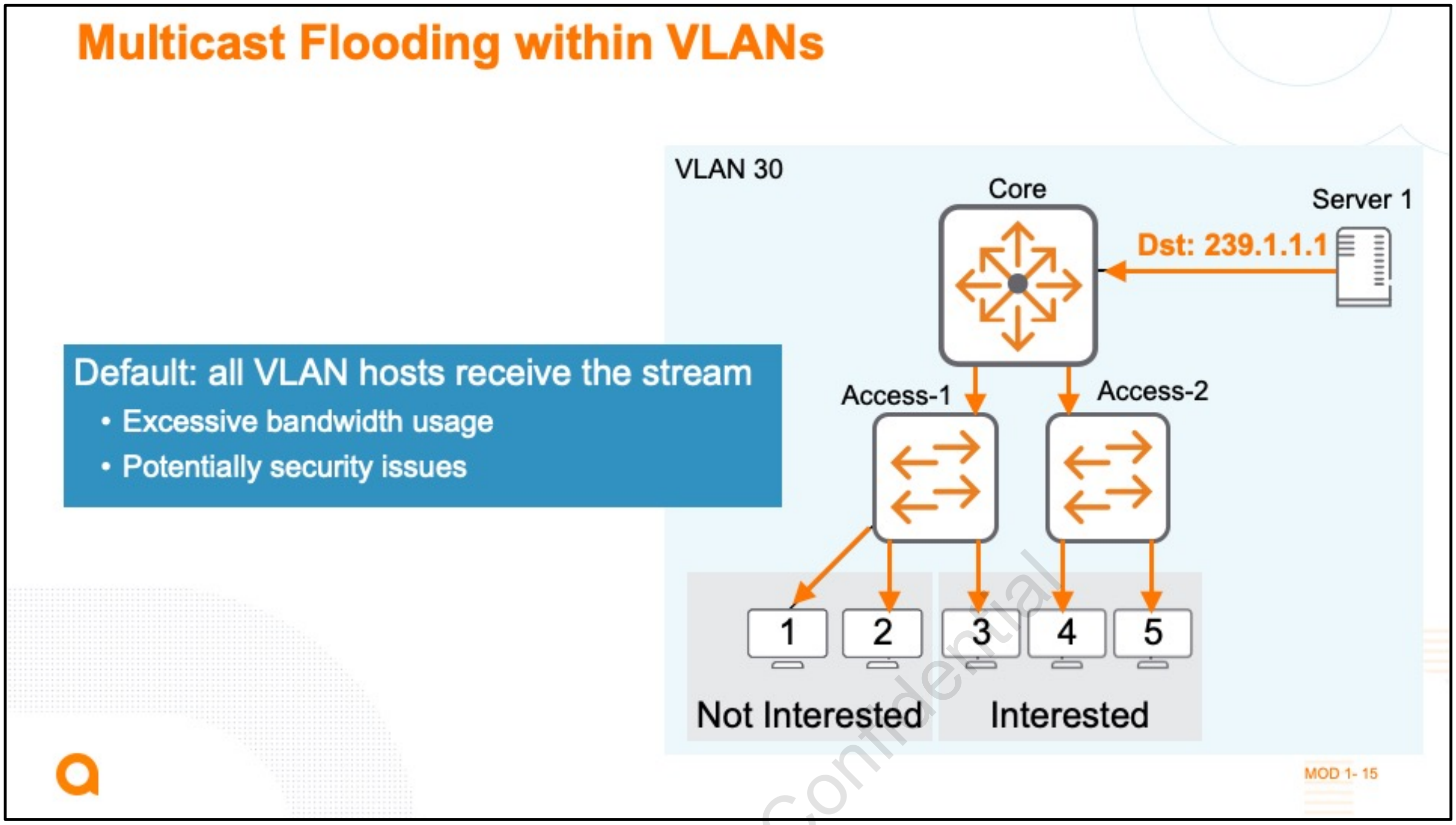
IGMP manages multicast group memberships based on a query and response mechanism. The multicast group uses four fundamental types of messages to communicate:

Query: A message sent from the querier asking for a response from each host belonging to the multicast group. If a multicast router supporting IGMP is not present, the switch must assume this function to elicit group membership information from the hosts on the network..

Join: A message sent by a host to the querier to indicate that the host wants to be or is a member of a given group indicated in the join message.

Leave: A message sent by a host to the querier to indicate that the host has ceased to be a member of a specific multicast group.

Report: A message sent by a host, in response to a multicast router query, identifying the particular multicast stream or streams the host wishes to receive.



Multicast Flooding within VLANs

Default: all VLAN hosts receive the stream

- Excessive bandwidth usage
- Potentially security issues

By default, Ethernet switches flood multicasts on all ports in a VLAN, or broadcast domain. Thus, although multicasts should target only the specific devices that require them, they act like broadcasts, creating congestion for an entire subnet. This can waste bandwidth and other resources. If all edge ports must carry traffic for all multicast groups within the VLAN, overall performance quickly suffers.

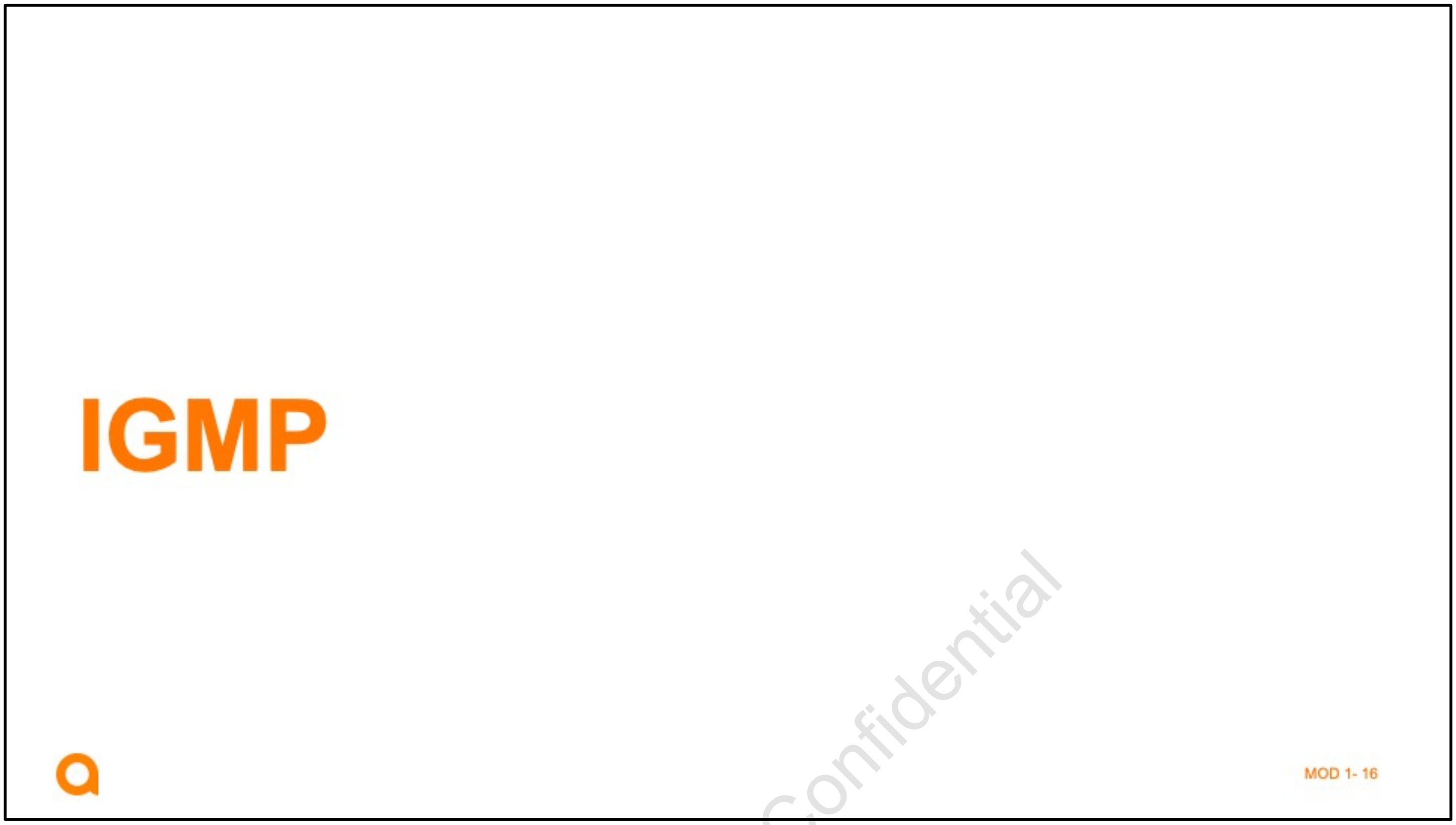
Flooding multicast packets can also increase security risk - every endpoint in a subnet receives traffic that is only meant for a particular group. Even if most users would not know how to eavesdrop on the multicast traffic from a neighboring meeting room's video conference that shows up unnecessarily on their ports, some users might—which could violate privacy regulations. Finally, service providers that provide for-pay services must limit multicast traffic to the properly registered receivers so that they can track who actually uses the services.

In summary, to deploy multicasting applications effectively, you must enable the switches to suppress specific multicasts on ports that do not require them.

Applications and processing

Applications are defined to use Layer-3 IP multicast addresses, but the actual hardware, like NICs or switches, are processing the Layer-2 addresses. By default, devices treat multicasts like broadcasts (with few exceptions). However, once multicasting support is enabled, devices can typically intelligently process the multicast traffic. For example, if a device did not have multicast enabled, the NIC would treat it as a broadcast and forward it up the protocol stack for the CPU to process. However, with multicasting enable, as shown in the figure, the NIC listens for only for multicast addresses that are relevant for upper-layer applications.

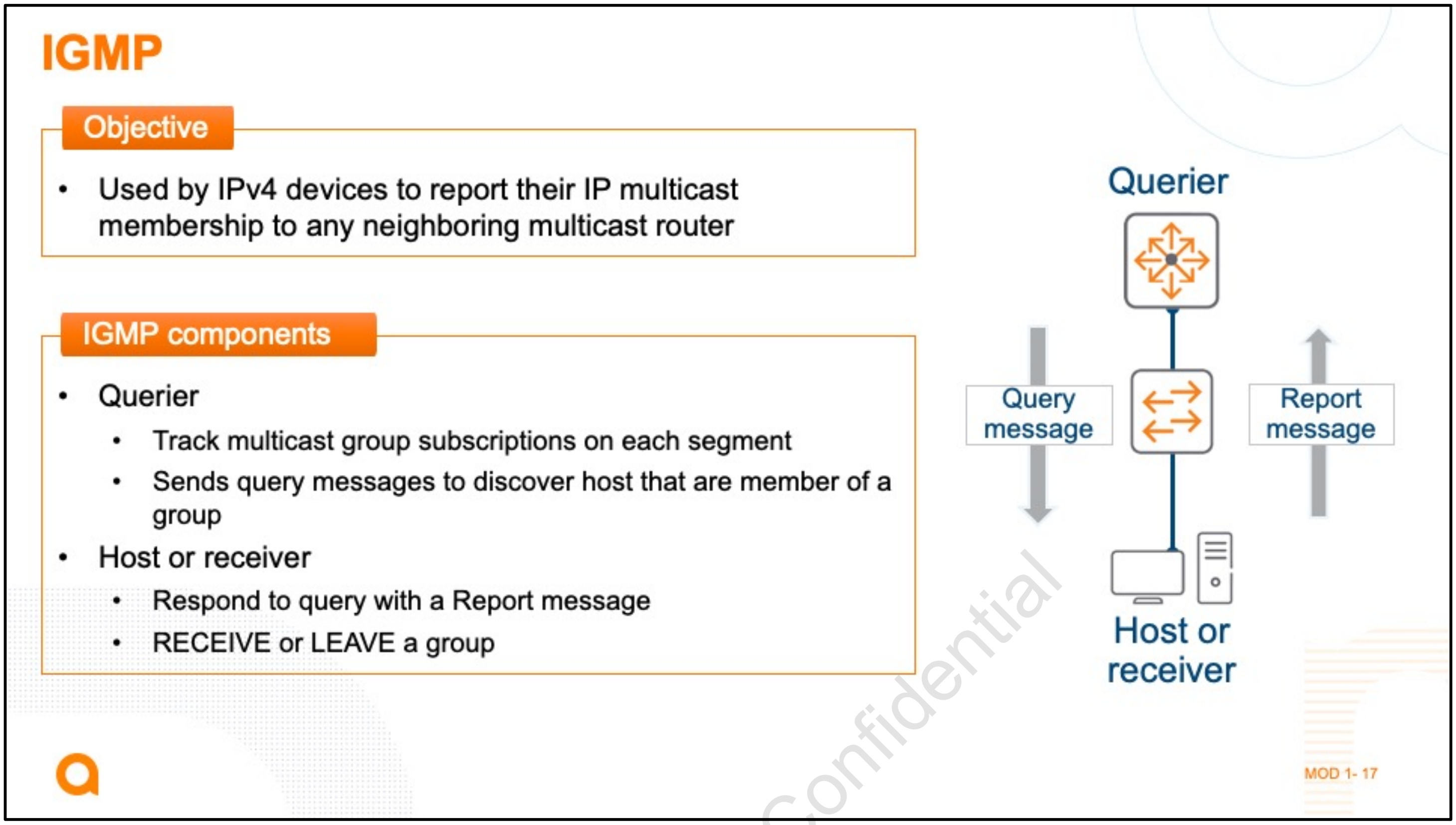
The second problem is that switches flood multicast traffic. In the figure, Client 1 and 2 receive multicast streams, even though they aren’t running a multicast application. This wastes CPU cycles on these clients to basically process and drop the multicast traffic. The switches need an intelligent method of forwarding/filtering multicasts instead of treating them as a broadcast.



IGMP

MOD 1- 16

Let’s dive into the IGMP details.



IGMP

Objective

- Used by IPv4 devices to report their IP multicast membership to any neighboring multicast router

IGMP components

- Querier
  - Track multicast group subscriptions on each segment
  - Sends query messages to discover host that are member of a group
- Host or receiver
  - Respond to query with a Report message
  - RECEIVE or LEAVE a group

MOD 1- 17

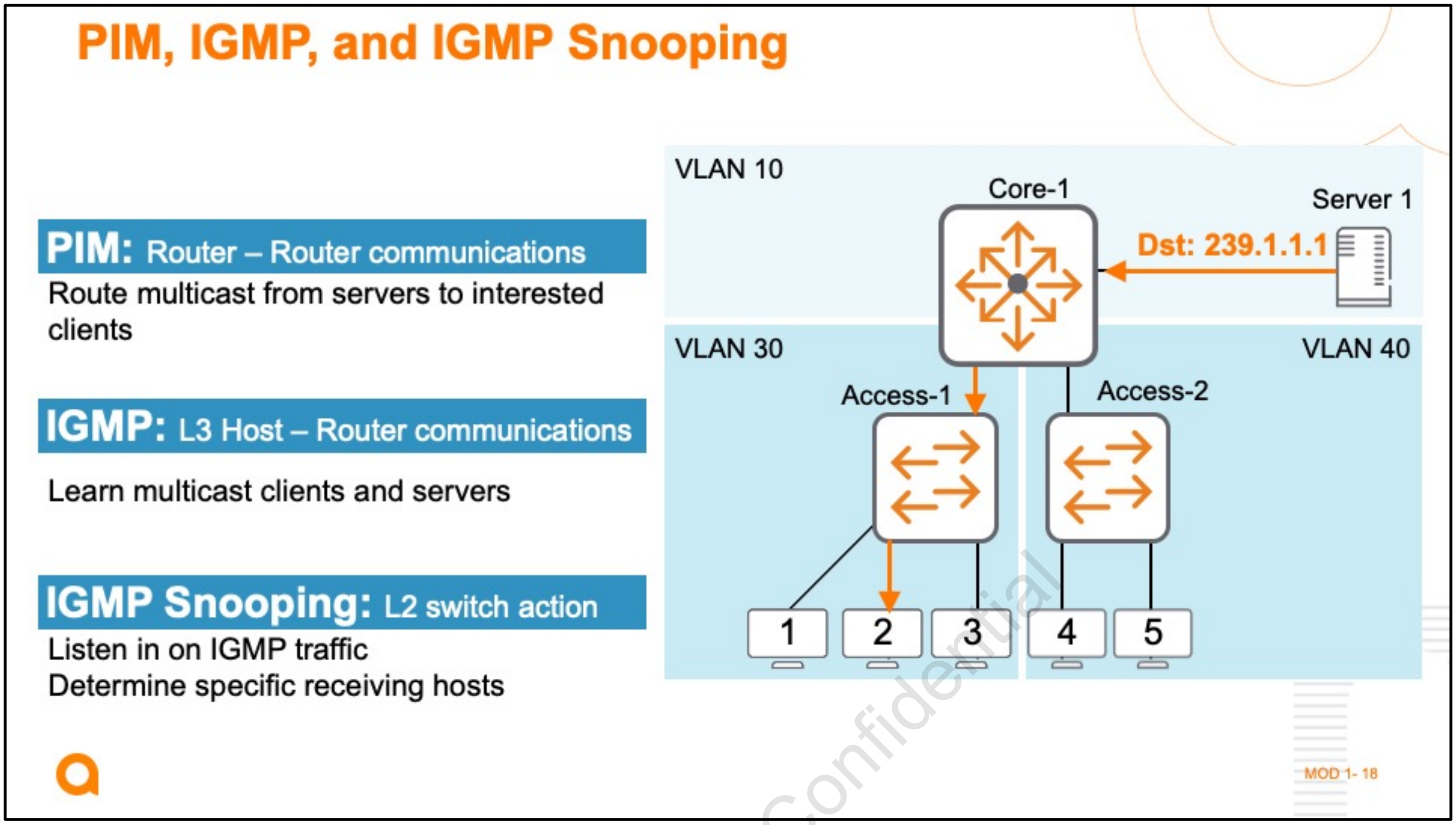
Multicast protocols are focused on “group membership”. This simply indicates which multicast streams are desired. If a host wishes to receive a multicast stream destined to address 239.1.1.1, they join group 239.1.1.1.

IPv4 systems (hosts and routers) use IGMP to report their IP multicast group memberships to any neighboring multicast routers. IGMP messages are encapsulated in IP datagrams, with IP protocol number 2.

IGMP establishes two device types - Querier and Hosts or receivers

A querier is a Layer 3 switch that runs IGMP. It sends periodic queries, asking hosts in the subnet which multicast group traffic they wish to receive. A multicast group is commonly referred to as a multicast stream - the application information, like a video stream.

A multicast receiver is a host that wishes to receive a multicast stream or streams, like a video conference unit or a user’s desktop.



PIM, IGMP, and IGMP Snooping

PIM: Router – Router communications

Route multicast from servers to interested clients

IGMP: L3 Host – Router communications

Learn multicast clients and servers

IGMP Snooping: L2 switch action

Listen in on IGMP traffic

Determine specific receiving hosts

MOD 1- 18

Three mechanisms work as a team to intelligently deliver multicast streams to interested clients, and only interested clients. They maximize multicast efficiency by minimizing multicast overhead.

Protocol Independent Multicast (PIM) is a router-to-router communications protocol – used by routers to build a forwarding tree and route traffic from the source toward interested clients. You will learn about PIM in the next module of this course.

The Internet Group Message Protocol (IGMP) is a host-to-router communications protocol – used by IPv4 devices to report their IP multicast membership to any neighboring multicast router. It helps to determine who is interested in receiving a multicast stream. Because of IGMP, router Core-1 knows that zero clients in VLAN 40 are interested in receiving the multicast stream, and more than zero clients in VLAN 30 are interested in the stream. Thus, Core-1 only forwards the stream out its VLAN 30 interface – no wasted bandwidth on VLAN 40!

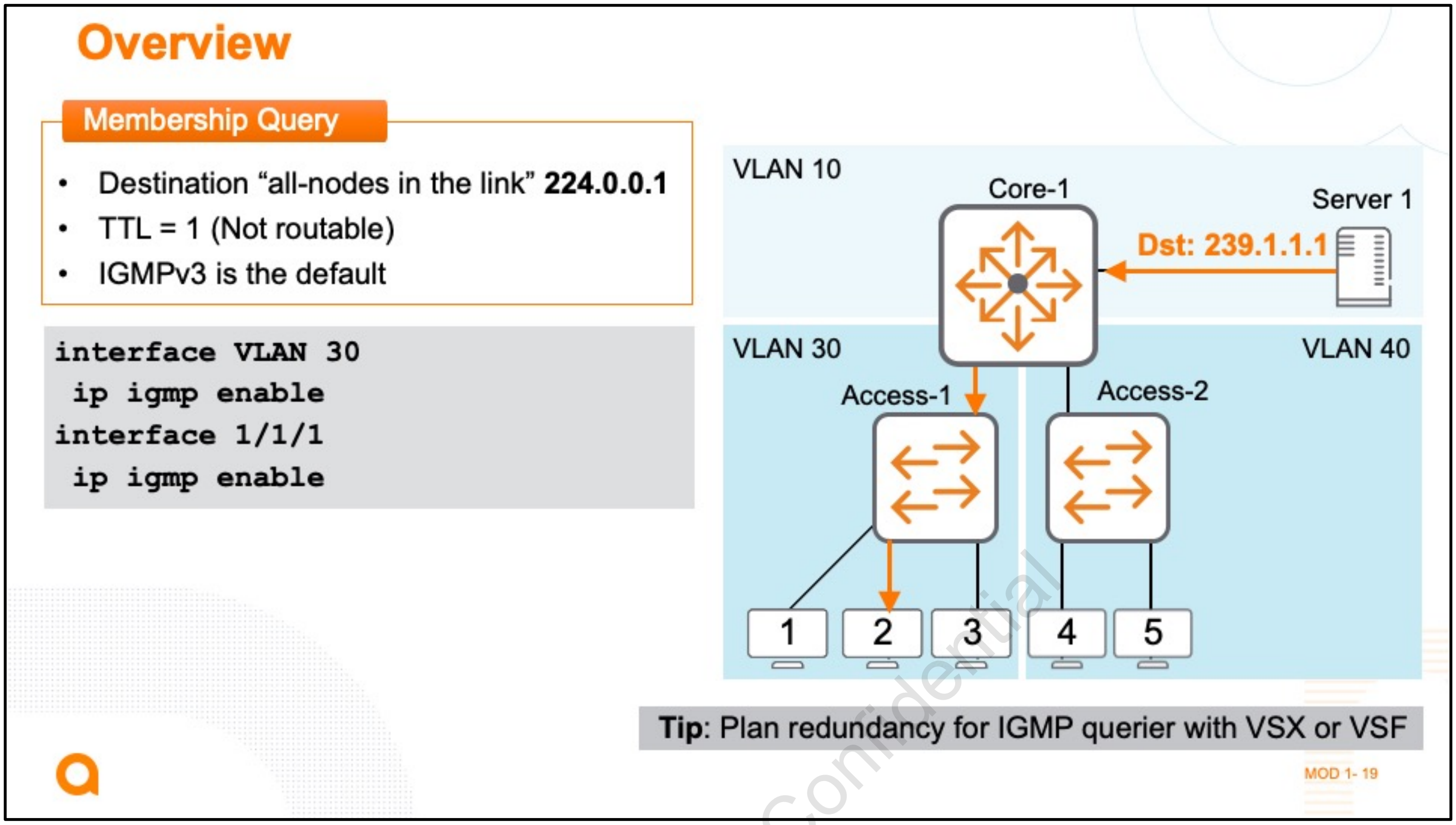
However, by default L2 switches flood all Broadcast, Unknown unicast, and Multicast (BUM) traffic out all ports in the VLAN (other than the ingress port). This means that hosts 1 – 3 would receive the stream, even though only host 2 is interested. When you enable IGMP snooping, Layer 2 switches listen in (snoop) on IGMP traffic between hosts and routers. They learn the MAC addresses of hosts that are actually interested in the stream and modify their default behavior accordingly. They only forward the stream to interested hosts in the VLAN – no wasted bandwidth in VLAN 30!

Notice that PIM and IGMP are protocols, while IGMP snooping is simply a mechanism that you enable on switches. The focus of this module is the IGMP protocol and IGMP snooping.

Note:

The IGMP functions fulfilled by switches are sometimes called IGMP snooping because switches traditionally did not generate IGMP messages, but only listened for the ones sent by routers and hosts.

However, routing AOS-CX switches can fulfill all IGMP functions. They can run the IGMP protocol as a Layer 3 router, and perform IGMP snooping as Layer 2 switch. You enable IGMP on a routed. L3 interface, while you enable IGMP snooping on a Layer 2 VLAN. Whether the VLAN has an IP address has no effect on this snooping functionality.



Overview

Membership Query

- Destination “all-nodes in the link” **224.0.0.1**
- TTL = 1 (Not routable)
- IGMPv3 is the default

```
interface VLAN 30
 ip igmp enable
interface 1/1/1
 ip igmp enable
```

Tip: Plan redundancy for IGMP querier with VSX or VSF

MOD 1- 19

To determine group membership, the querier (router) sends a message to every host on the subnet. In the figure, router Core-1 is the IGMP querier - Access 1 and 2 are Layer-2 switches. The querier must maintain a list of hosts in the subnet interested in multicast flows. To do this, the query message is multicast to the “all-hosts” multicast address of 224.0.0.1. This message is created with a TTL of 1 which indicates its non-routable nature.

Generally, the default router for a VLAN should act as the IGMP querier. The figure shows how to enable IGMP on an AOS-CX routing switch, both for a VLAN interface and for a physical Layer 3 interface. IGMPv3 is the default implementation. Use the disable parameter to disable IGMP on a Layer 3 interface. Note that if you disable IGMP on an interface and then later re-enable IGMP on that VLAN, the switch restores the last-saved IGMP configuration for that interface. Use the no ip igmp command to completely remove it from an interface.

Tip

You should plan redundancy for the IGMP querier role by implementing either VSX or VSF on your AOS-CX switches.

Multiple IGMP queriers and elections

Just as you need to consider redundancy for the default gateway, you should consider how you will provide redundancy for the IGMP querier role. In the example shown above, the core is a VSF fabric, which has built-in redundancy.

If you are using a protocol such as Virtual Router Redundancy Protocol (VRRP), you should usually implement a backup IGMP querier. Simply enable the querier function on both VRRP devices. Even if you do not have redundancy for routing within the VLAN, though, you can still create a backup IGMP querier. Just add an IP address on the VLAN to one of the non-routing switches and enable IGMP.

Only one device acts as the IGMP querier within the VLAN. Other querier-capable switches become IGMP listeners. As you will see, a listener continues to use IGMP to learn the multicast traffic that endpoints desire in the interface, just like the querier. It simply refrains from sending queries itself. The listener also takes over as querier if it fails to hear a query within two times the querier interval.

The AOS-Switch behavior in determining whether to act as IGMP querier or listener differs based on whether the switch is also implementing a multicast routing protocol such as PIM. You will learn about PIM in the next module, but the behavior with PIM enabled is described below for your reference.

Behavior without PIM enabled

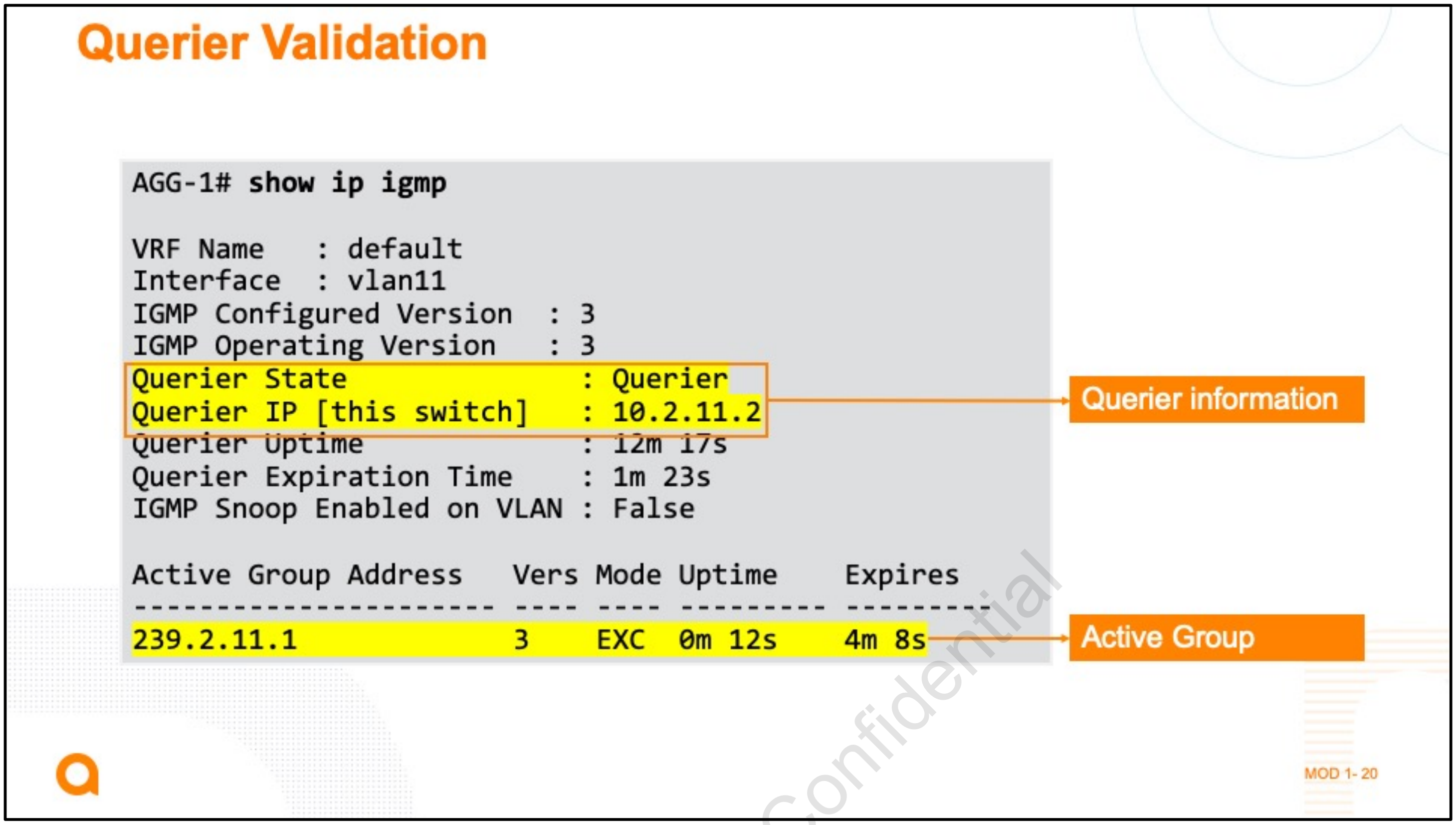
When you enable IGMP on a switch VLAN with an IP address, the querier capability is enabled by default. However, based on the presumption that a multicast router will act as IGMP querier, the switch does not attempt to become querier immediately. Instead it listens for IGMP queries for a waiting period equal to twice the IGMP querier interval. If it hears any queries in this period, it becomes an IGMP listener. If it does not, it becomes the querier. However, if the switch later hears a query from another device, it disables the querier function and defers to that device.

If you enable IGMP on two AOS-Switches that are not running PIM, the first switch on which you enable IGMP becomes the querier.

Behavior with PIM enabled

When an AOS-Switch has PIM and IGMP enabled on a VLAN, it uses the IGMPv2 rules for its querier functions. As soon as IGMP is enabled, the switch begins to send IGMP queries at the designated interval. If it ever hears a query from another querier with a lower IP address, it stops sending queries and becomes an IGMP listener. In other words, the switch running IGMP and PIM with the lowest IP address becomes querier, possibly pre-empting the role from another switch.

IGMP operates separately from Virtual Router Redundancy Protocol (VRRP). That is, if two routing switches implement VRRP and IGMP in VLAN 20, either the VRRP backup or the VRRP master could be elected as the IGMP querier regardless of its VRRP role. This is because the switches use their actual IP addresses for the election (the VRRP master does not use its virtual address).



Querier Validation

```
AGG-1# show ip igmp

VRF Name    : default
Interface   : vlan11
IGMP Configured Version  : 3
IGMP Operating Version   : 3
Querier State                : Querier
Querier IP [this switch]     : 10.2.11.2
Querier Uptime               : 12m 17s
Querier Expiration Time      : 1m 23s
IGMP Snoop Enabled on VLAN : False

Active Group Address   Vers Mode Uptime    Expires
---------------------- ---- ---- --------- ---------
239.2.11.1             3    EXC  0m 12s    4m 8s
```

The show ip igmp command displays information about the querier. Notice that the output includes which device is running the querier role for the segment.

This command includes information such as:

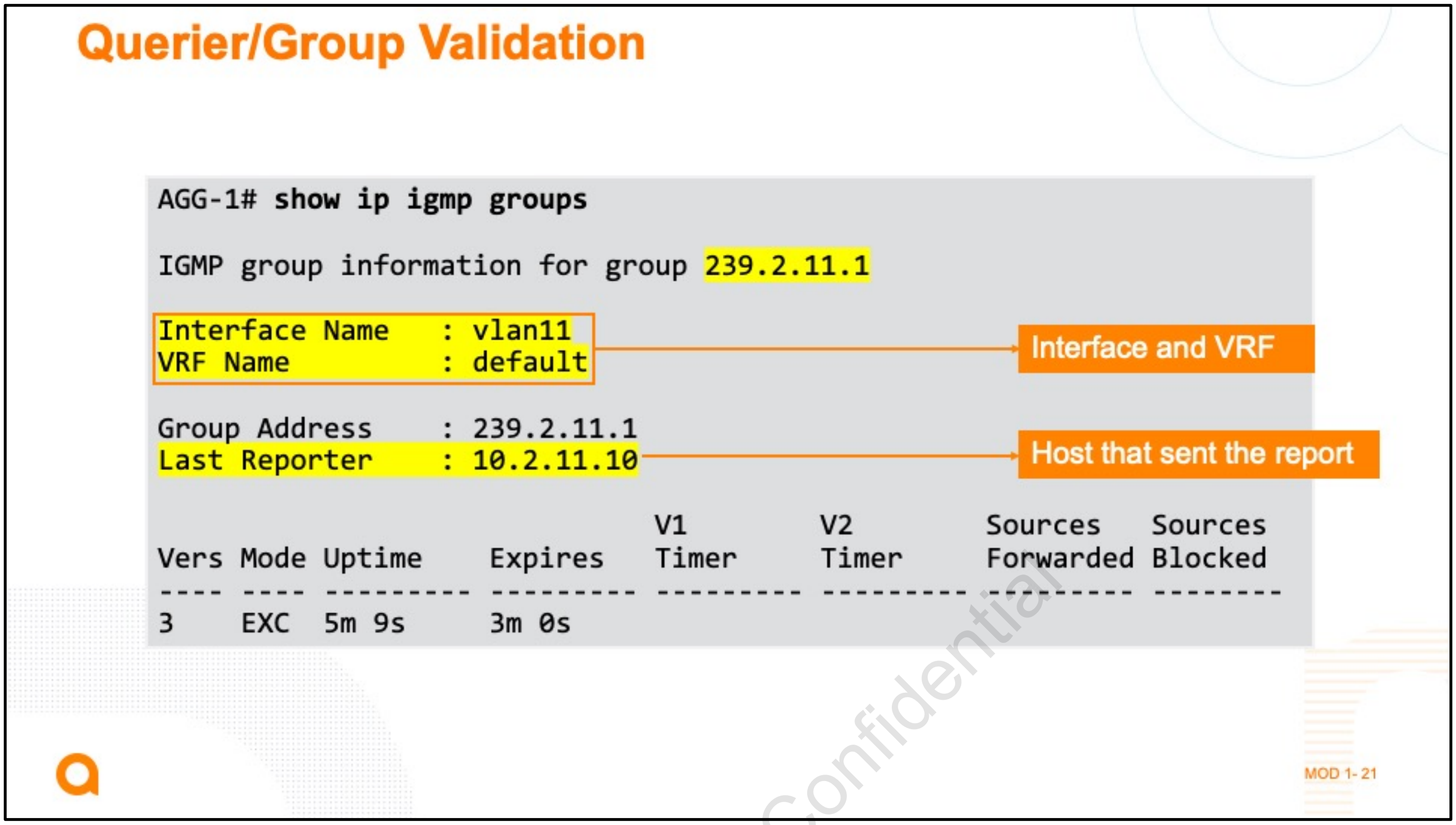
VRF Name

Layer 3 Interface

IGMP Configured Version (v3 by default)

Querier State

Querier IP address



Querier/Group Validation

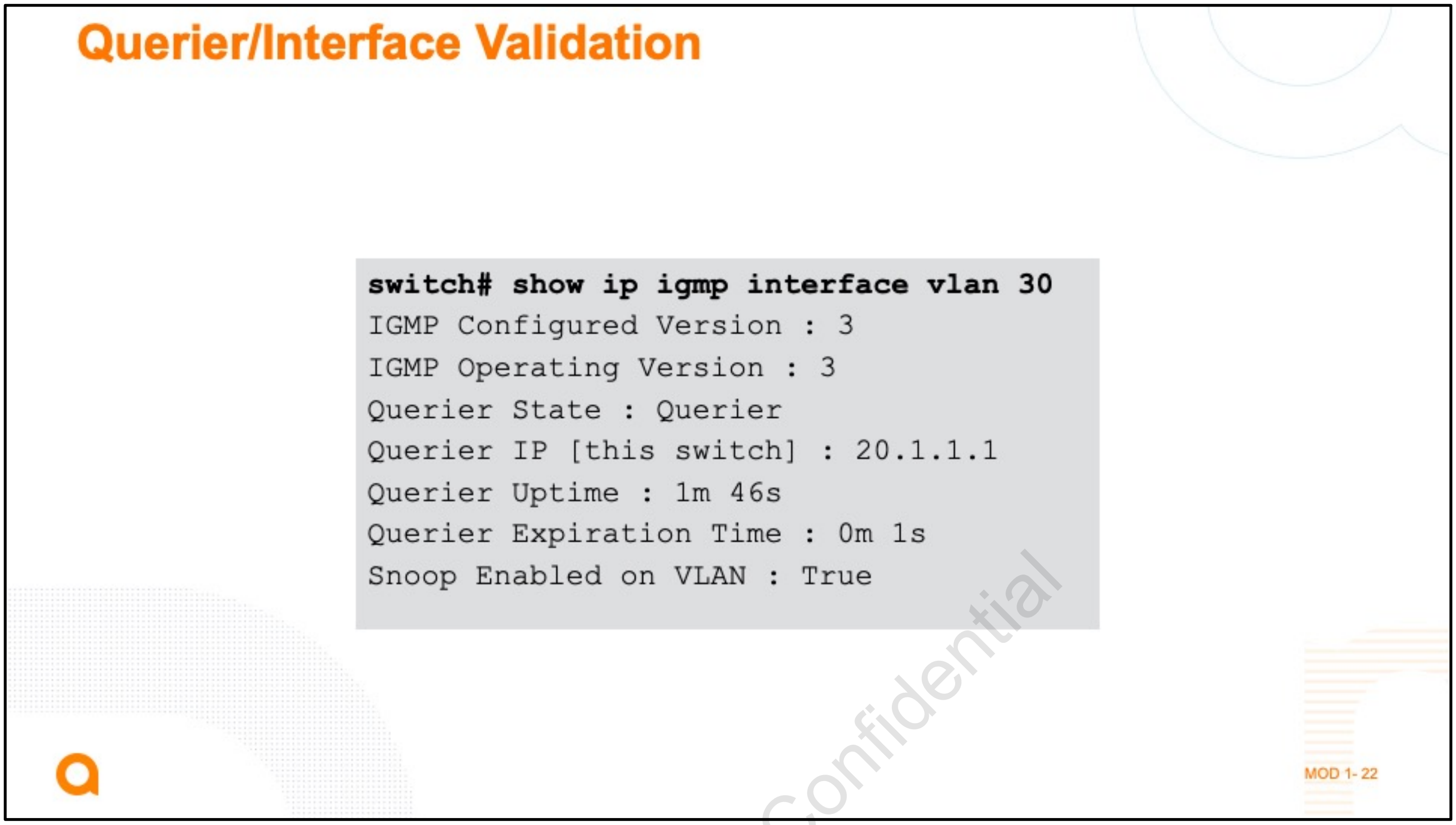
```
AGG-1# show ip igmp groups

IGMP group information for group 239.2.11.1

Interface Name    : vlan11
VRF Name          : default

Group Address     : 239.2.11.1
Last Reporter     : 10.2.11.10

                                V1        V2        Sources   Sources
Vers Mode Uptime    Expires   Timer     Timer     Forwarded Blocked
---- ---- --------- --------- --------- --------- --------- --------
3    EXC  5m 9s     3m 0s
```

Interface and VRF

Host that sent the report

MOD 1- 21

The show ip igmp groups comand display information about the multicast groups including the interface and the VRF from where the switch receive these packets.

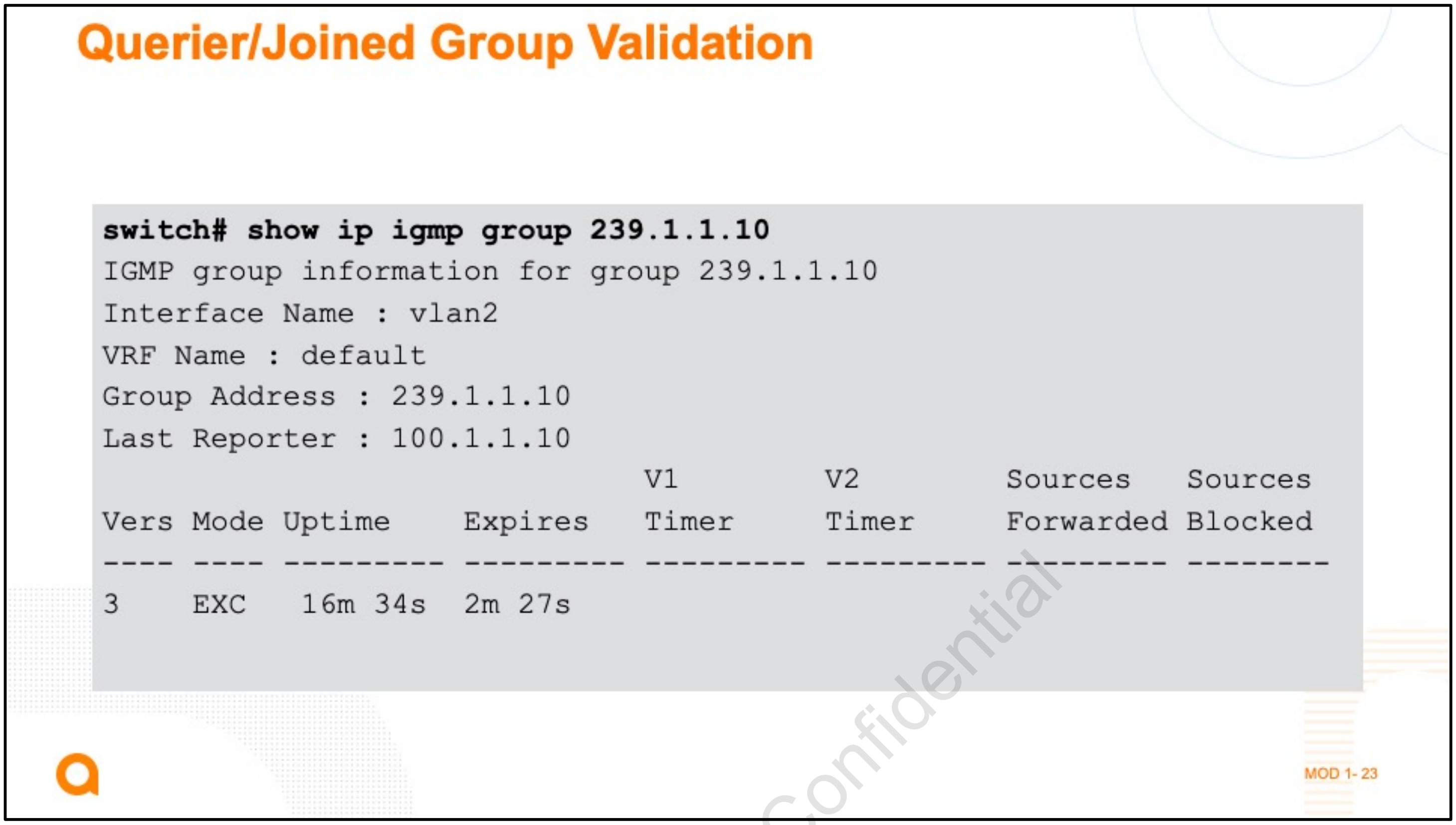


Querier/Interface Validation

```
switch# show ip igmp interface vlan 30
IGMP Configured Version : 3
IGMP Operating Version : 3
Querier State : Querier
Querier IP [this switch] : 20.1.1.1
Querier Uptime : 1m 46s
Querier Expiration Time : 0m 1s
Snoop Enabled on VLAN : True
```

MOD 1- 22

You could also use the command show ip igmp interface to see IGMP interface-specific configuration and operation info.



Querier/Joined Group Validation

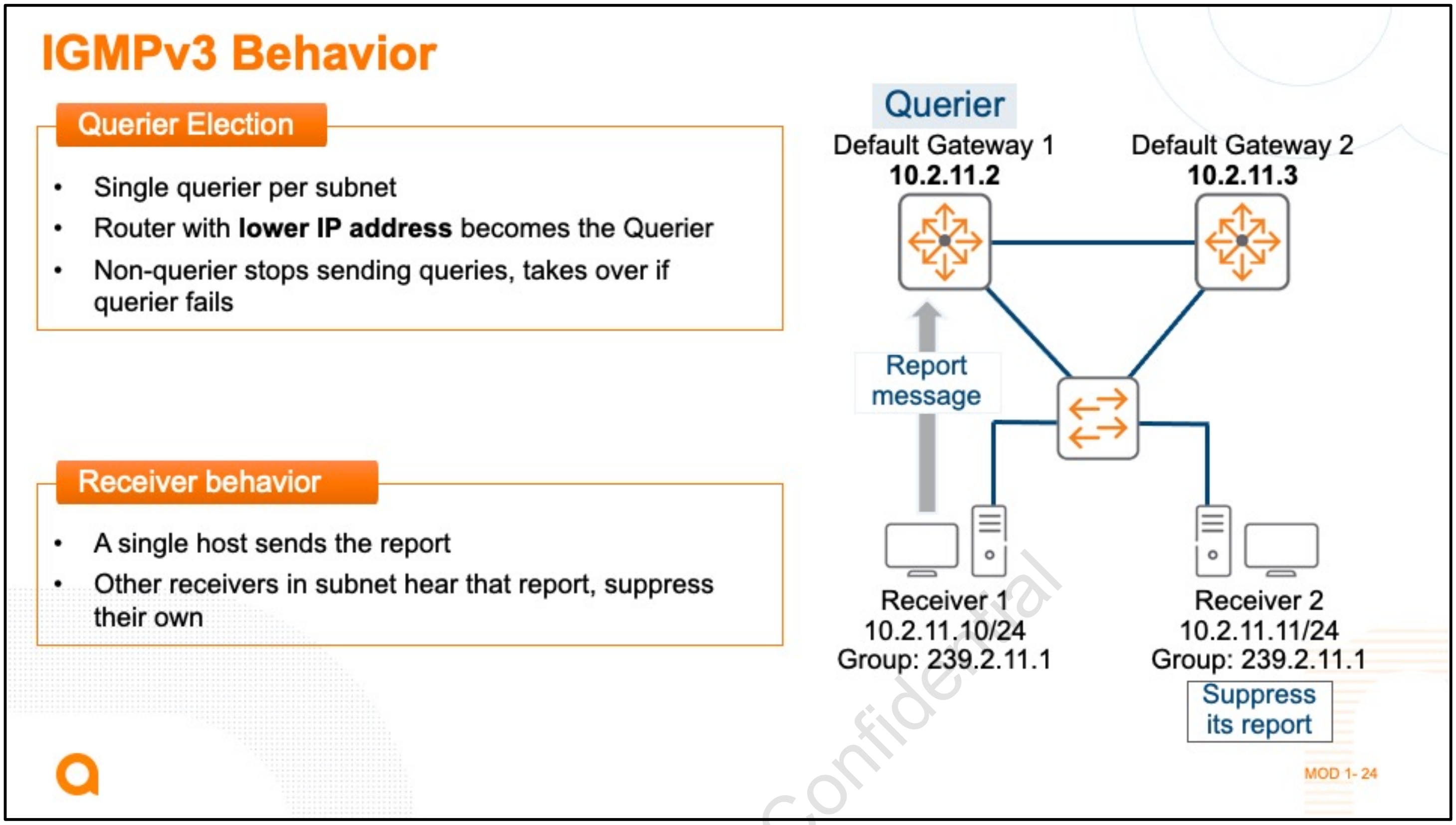
```
switch# show ip igmp group 239.1.1.10
IGMP group information for group 239.1.1.10
Interface Name : vlan2
VRF Name : default
Group Address : 239.1.1.10
Last Reporter : 100.1.1.10
                              V1        V2        Sources   Sources
Vers Mode Uptime    Expires   Timer     Timer     Forwarded Blocked
---- ---- --------- --------- --------- --------- --------- --------
3    EXC   16m 34s  2m 27s
```

MOD 1- 23

Validate IGMP joined group information with this command, as shown in the figure:

Switch# show ip igmp group <GROUP-IP> [source <SOURCE-IP>] [vrf <VRF-NAME> | all-vrfs] [vsx-peer]

You can also use the show ip igmp groups command, which will list all IGMP groups the querier knows about. Many other show commands are available to verify the operation of IGMP.



IGMPv3 Behavior

Querier Election

- Single querier per subnet
- Router with **lower IP address** becomes the Querier
- Non-querier stops sending queries, takes over if querier fails

Receiver behavior

- A single host sends the report
- Other receivers in subnet hear that report, suppress their own

Just as you should consider redundancy for the default gateway, you should also provide redundancy for the IGMP querier role. IGMPv3 elects a single querier per subnet. Other querier-capable switches become IGMP listeners. A listener continues to use IGMP to learn the multicast traffic that endpoints desire in the interface, just like the querier. It simply refrains from sending queries itself. The listener also takes over as querier if it does not hear a query within two times the querier interval.

When a host receives an IGMP query packet, it kicks off a timer that begins with a random value that is less than the Maximum Response Time. If no other host responds with a membership report before this random timer expires, the host will then reply with a report. This decreases the number of total IGMP reports needed to maintain the group state. This preserves local bandwidth, because the host suppresses its own reports unless absolutely necessary.

The AOS-Switch behavior in determining whether to act as IGMP querier or listener differs based on whether the switch is also implementing a multicast routing protocol such as PIM. You will learn about PIM in the next module, but the behavior with PIM enabled is described below for your reference.

Behavior without PIM enabled

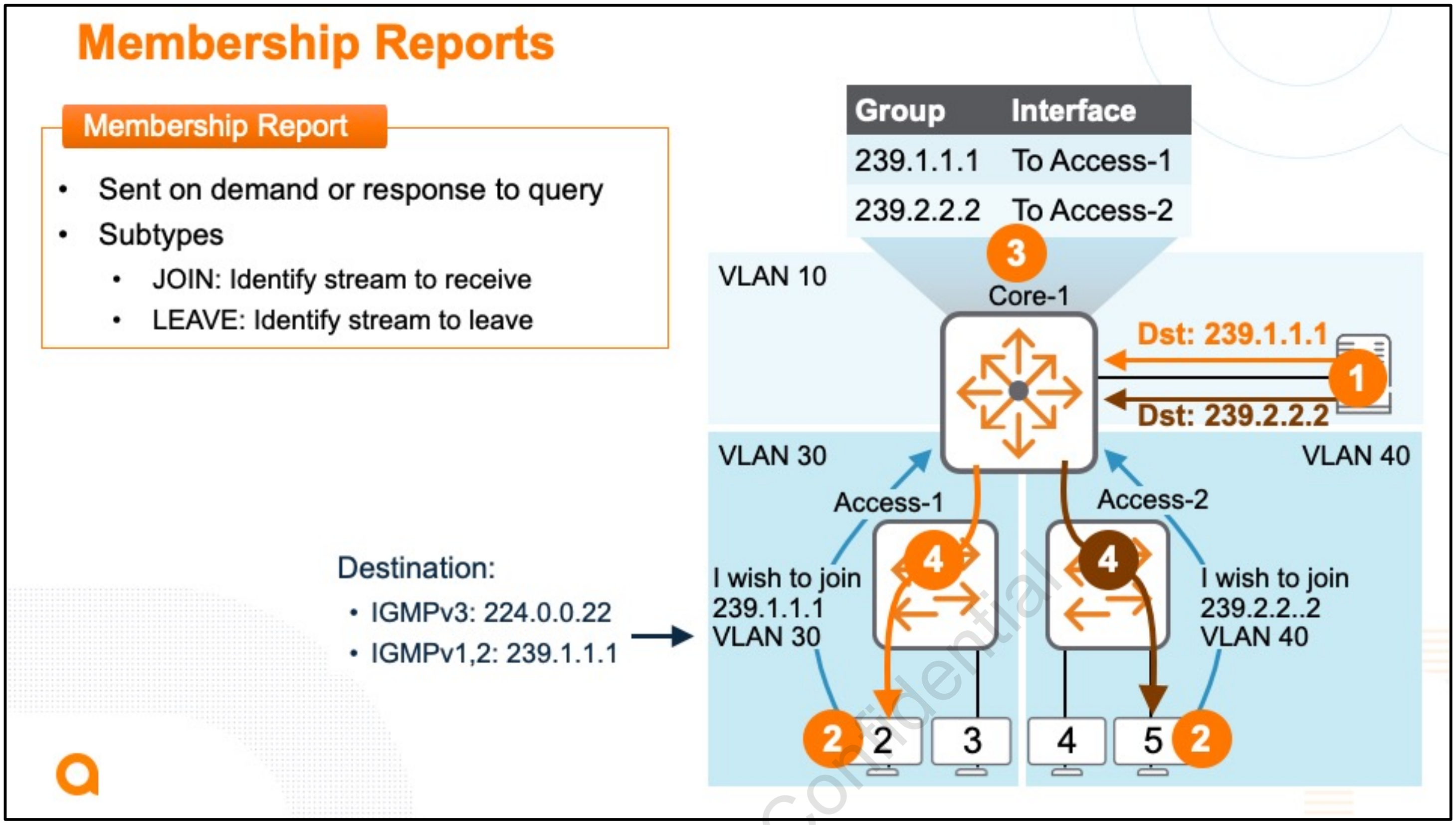
When you enable IGMP on a switch VLAN with an IP address, the querier capability is enabled by default. However, based on the presumption that a multicast router will act as IGMP querier, the switch does not attempt to become querier immediately. Instead it listens for IGMP queries for a waiting period equal to twice the IGMP querier interval. If it hears any queries in this period, it becomes an IGMP listener. If it does not, it becomes the querier. However, if the switch later hears a query from another device, it disables the querier function and defers to that device.

If you enable IGMP on two AOS-Switches that are not running PIM, the first switch on which you enable IGMP becomes the querier.

Behavior with PIM enabled

When an AOS-Switch has PIM and IGMP enabled on a VLAN, it uses the IGMPv2 rules for its querier functions. As soon as IGMP is enabled, the switch begins to send IGMP queries at the designated interval. If it ever hears a query from another querier with a lower IP address, it stops sending queries and becomes an IGMP listener. In other words, the switch running IGMP and PIM with the lowest IP address becomes querier, possibly pre-empting the role from another switch.

IGMP operates separately from Virtual Router Redundancy Protocol (VRRP). That is, if two routing switches implement VRRP and IGMP in VLAN 20, either the VRRP backup or the VRRP master could be elected as the IGMP querier regardless of its VRRP role. This is because the switches use their actual IP addresses for the election (the VRRP master does not use its virtual address).



Membership Reports

Membership Report

- Sent on demand or response to query
- Subtypes
  - JOIN: Identify stream to receive
  - LEAVE: Identify stream to leave

Destination:

- IGMPv3: 224.0.0.22
- IGMPv1,2: 239.1.1.1

Group	Interface
239.1.1.1	To Access-1
239.2.2.2	To Access-2

Membership report messages are sent by IP endpoints to report (to neighboring routers) interest in joining or leaving a multicast stream. These messages are sent either in response to a membership query message or when the multicast application in the endpoint starts. There are two subtypes of membership messages:

JOIN: Sent from host to querier to identify the multicast stream the host wishes to receive.

LEAVE: Sent from host to querier to indicate that the host shall longer be a member of a Multicast group.

In the figure, a server sends a multicast stream to 239.1.1.1 and another stream to 239.2.2.2. Depending on PIM configuration, these streams may not be forwarded anywhere – Core-1 discards the frames because nobody is interested in receiving the streams – there are no group members.

Then host 2 in VLAN 30 sends an IGMP group membership report, “ I wish to join 239.1.1.1”. Meanwhile, host 5 in VLAN 40 sends a group membership report, “I wish to join 239.2.2.2”. Assuming all devices run IGMPv3, the messages are sent to 224.0.0.22. For IGMPv1 and2, messages are sent to the group address – 239.1.1.1 or 239.2.2.2.

When Core-1 receives the first report for a particular multicast group on a VLAN interface, it adds a table for that group to the VLAN. Each VLAN that implements IGMP has its own forwarding tables. Because the switch also performs IGMP snooping at Layer 2, the physical interface on which the switch received the report becomes a forwarding interface for that group. If the switch receives another membership report on a different physical interface, it adds that interface to the forwarding table as well.

IGMPv1 group membership entries do not persist indefinitely. The switch also sets a timer for the group membership, which it resets every time that its query produces another report. If the timer expires, the switch removes the group membership from the interface.

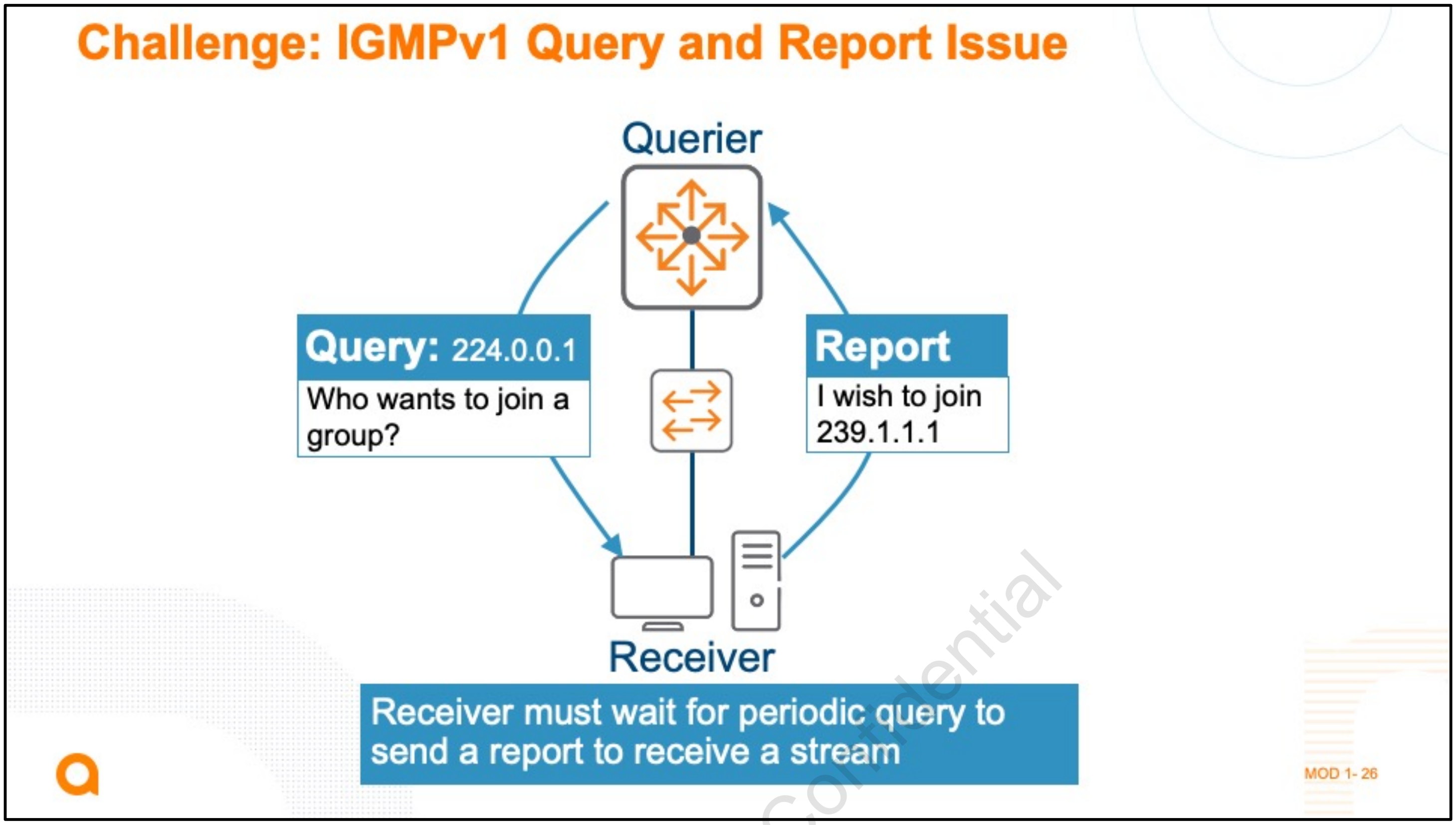
IGMPv2 and v3 have a faster way for an interface to leave a group. When a host no longer requires multicast traffic in a particular group, it sends an IGMP leave group message. The IGMP querier instantly sends one or more queries for the multicast group in question. This allows another IGMP host connected to the same interface to send a membership report indicating that it still wants to receive the multicasts.

--

Supplemental information on how AOS-CX switches handle reports

In the example above, only one IGMP host in VLAN 30 wants to join the 239.1.1.1 group. As you see, the core VSF fabric receives only one membership report. Even if multiple hosts on Access-1 wanted to join this group, though, they would all receive the membership report from the first host to submit the report because this report is a multicast. Those hosts would then suppress their own reports, and a single report still flows to the core. (This behavior is a bit different when Access-1 also implements IGMP, as you will see later.)

IGMP defines this behavior because it was originally designed to inform multicast routers if at least one host in the subnet requires multicasts in the subnet. How many hosts required the multicasts does not really matter. Because routing AOS-CX switches also implement IGMP at Layer 2, and need to discover on which physical interfaces to forward traffic, they alter this behavior a bit. The core VSF switch, which implements IGMP, suppresses the membership reports that it receives on each physical interface and does not forward them on other physical interfaces in the VLAN. This behavior prevents hosts on Access-2 from hearing the membership reports from hosts on Access-1. In this way, the core fabric can receive a report on the Access-2 link—for example, if any hosts connected to that switch wanted to listen on 239.1.1.1.



Challenge: IGMPv1 Query and Report Issue

Querier

Query: 224.0.0.1
Who wants to join a group?

Report
I wish to join 239.1.1.1

Receiver

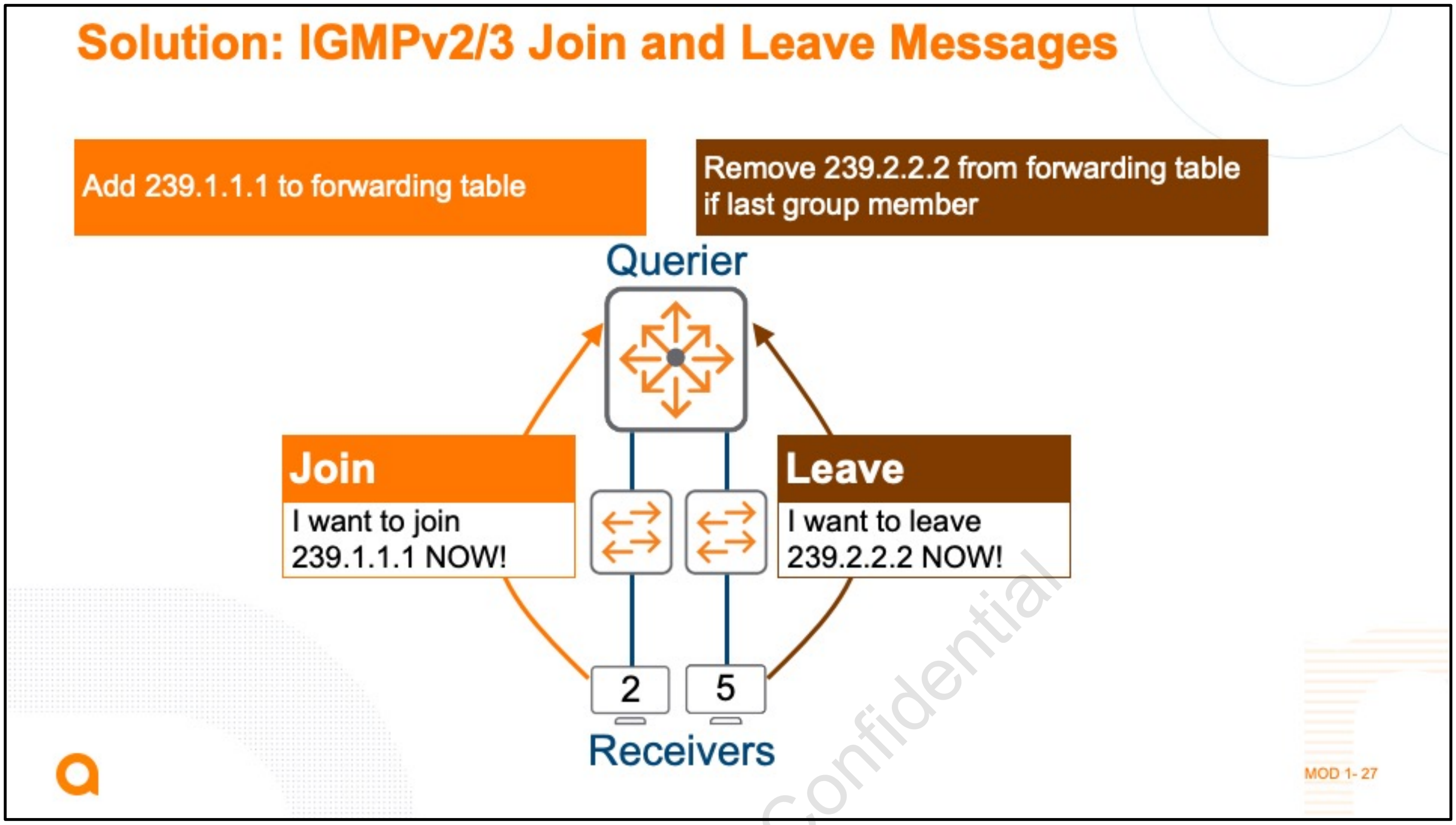
Receiver must wait for periodic query to send a report to receive a stream

MOD 1- 26

The original IGMPv1 only supported two message types: queries and reports. There are two problems with the original implementation and these messages types:

The host has to wait for a query before sending a report; this slows down access to the stream the in which the user is interested.

When a user is done with the multicast application, the host would still receive the stream until a corresponding query would resolve the issue, wasting bandwidth.



Solution: IGMPv2/3 Join and Leave Messages

Add 239.1.1.1 to forwarding table

Remove 239.2.2.2 from forwarding table if last group member

Querier

Join

I want to join 239.1.1.1 NOW!

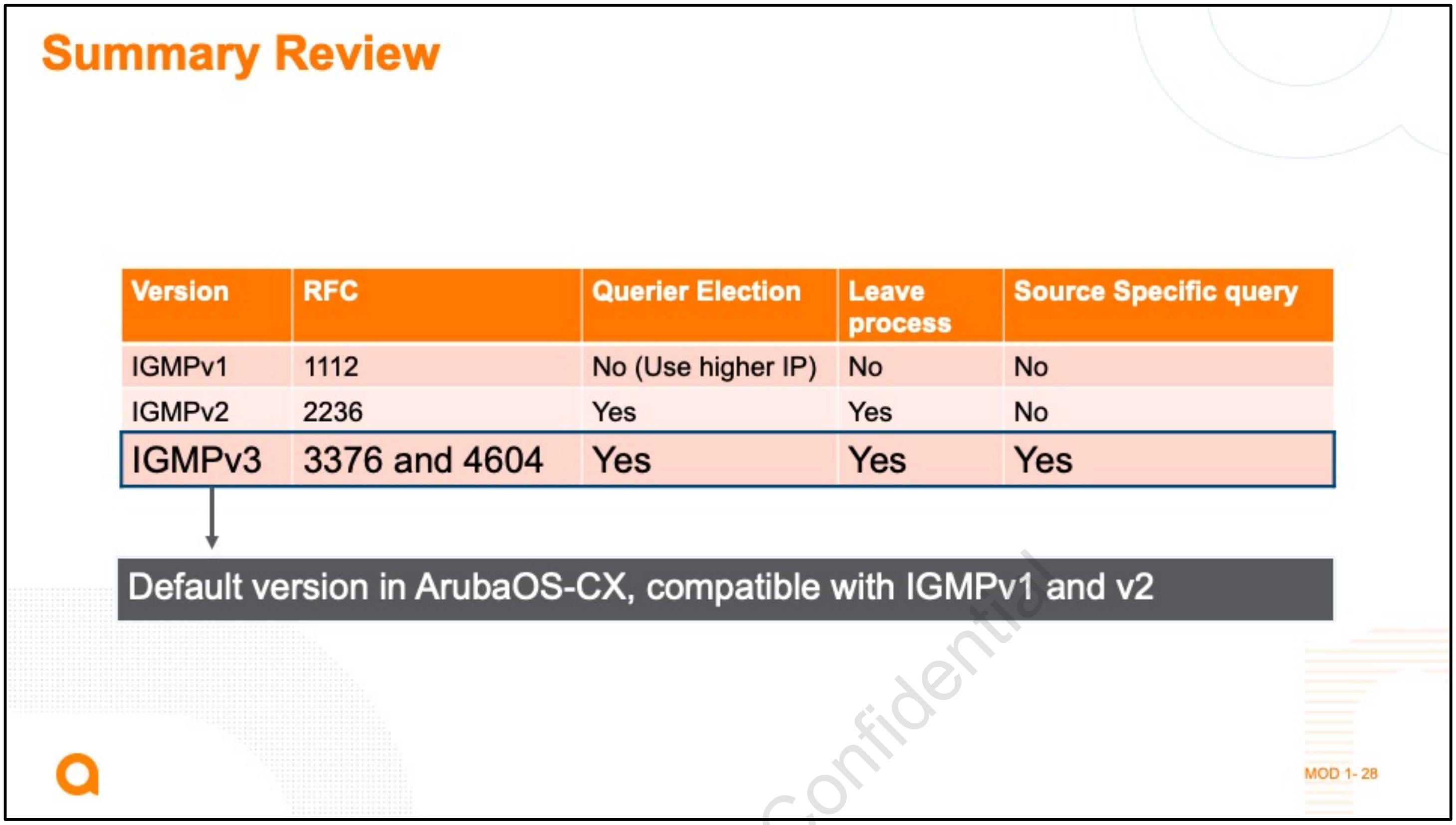
Leave

I want to leave 239.2.2.2 NOW!

Receivers

MOD 1- 27

IGMPv2 and v3 introduce two new message types: joins and leaves. A join message allows a multicast host to immediately signal a querier of a multicast group that the host needs to receive, expediting the receipt of the stream. And a leave message allows a host to immediately notify a querier when the host no longer needs the stream for the multicast group.



Summary Review

Version	RFC	Querier Election	Leave process	Source Specific query
IGMPv1	1112	No (Use higher IP)	No	No
IGMPv2	2236	Yes	Yes	No
IGMPv3	3376 and 4604	Yes	Yes	Yes

Default version in ArubaOS-CX, compatible with IGMPv1 and v2

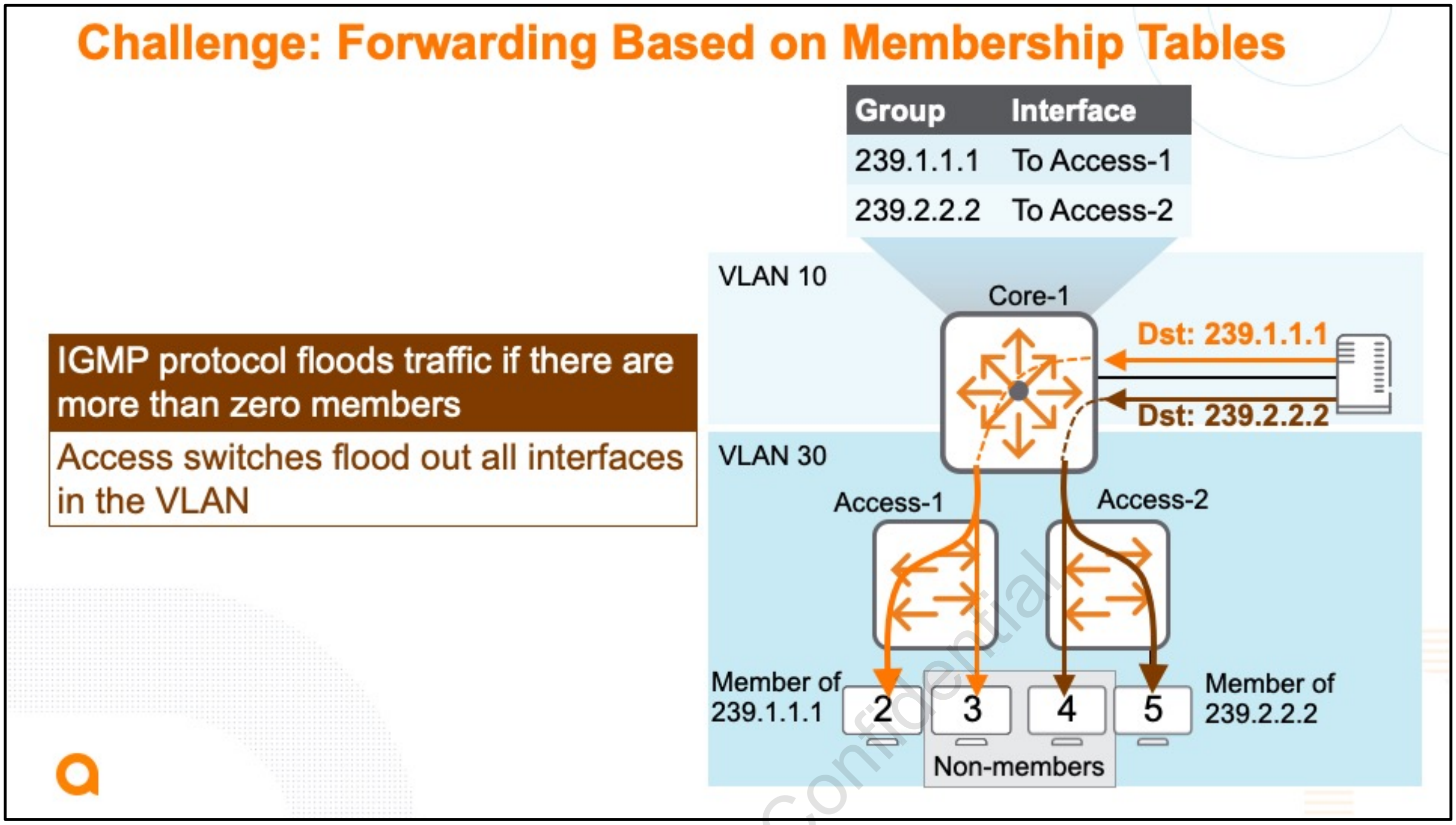
MOD 1- 28

The figure summarizes the differences between IGMP versions, as described below.

IGMPv1: Defined in RFC 1112, IGMPv1 offers a basic query-and-response mechanism to determine which multicast streams should be sent to a particular network segment. IGMPv1 has no mechanism for a host to signal that it wants to leave a group. When a host using IGMPv1 leaves a group, the router will continue to send the multicast stream until the group times out. As you can imagine, this can create a large amount of multicast traffic on a subnet if a host joins groups very quickly. IGMPv1 also does not elect a querier. If there are multiple queriers (routers) on the subnet, a designated router (DR) is elected using PIM to avoid sending duplicate multicast packets. The elected querier is the router with the highest IP address. IGMPv1 is rarely used in modern networks.

IGMPv2: Defined in RFC 2236, this newer version made improvements over IGMPv1. One of the most significant changes was the addition of a leave process. A host using IGMPv2 can send a leave-group message to the querier indicating that it is no longer interested in receiving a particular multicast stream. IGMPv2 also added group-specific-queries. This feature allows the querier to send a message to the host(s) belonging to a specific multicast group. The querier election process determines the querier based on a priority.

IGMPv3: Defined in RFC 3376 and 4604, this version 3 adds support for "source filtering“ - the ability for a system to report interest in receiving packets *only* from specific source addresses, or from *all but* specific source addresses, sent to a particular multicast address. Multicast routing protocols can use this information to avoid delivering multicast packets from specific sources to networks where there are no interested receivers.



Challenge: Forwarding Based on Membership Tables

Group	Interface
239.1.1.1	To Access-1
239.2.2.2	To Access-2

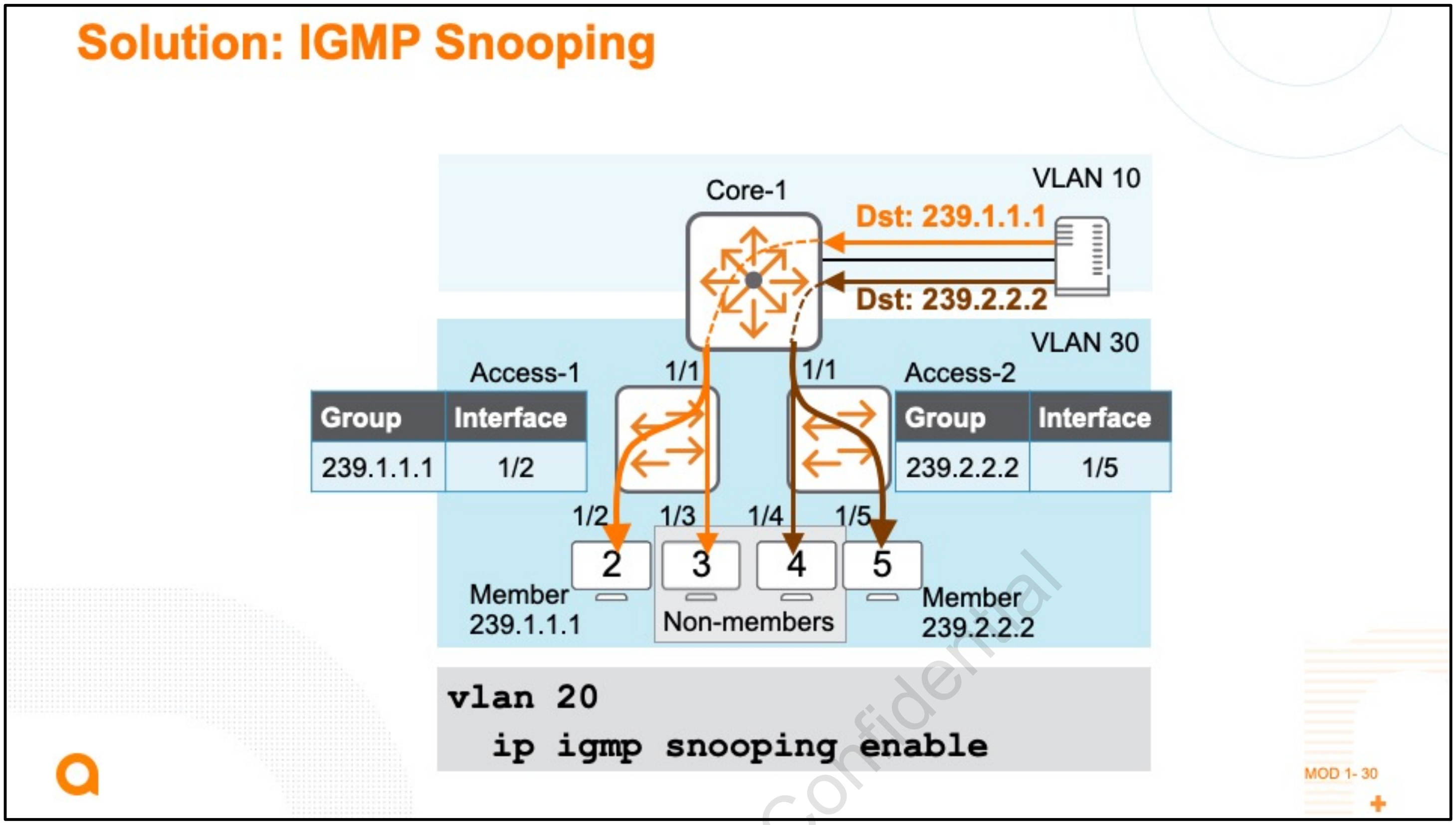
IGMP protocol floods traffic if there are more than zero members

Access switches flood out all interfaces in the VLAN

The multicast group table filters traffic for multicasts destined to that address. In this example, Core-1the core is the only switch running IGMP, and it is the IGMP querier. Based on the membership reports that it has received in VLAN 30, it has created two forwarding tables - one for 239.1.1.1 and one for 239.2.2.2. Bandwidth is saved – The 239.1.1.1 stream is only sent toward Access-1, and 239.2.2.2 is only sent toward Access-2

However, recall the default L2 switch behavior for BUM traffic - multicast traffic received by Access-1 and Access-2 is flooded out all ports in the VLAN – this includes the ports connected to hosts that are not a member of any group, and have no need for this potentially high-bandwidth video stream.

In essence, the core filters multicasts, but the access layer switches do not.



Solution: IGMP Snooping

Group	Interface
239.1.1.1	1/2

Group	Interface
239.2.2.2	1/5

```
vlan 20
  ip igmp snooping enable
```

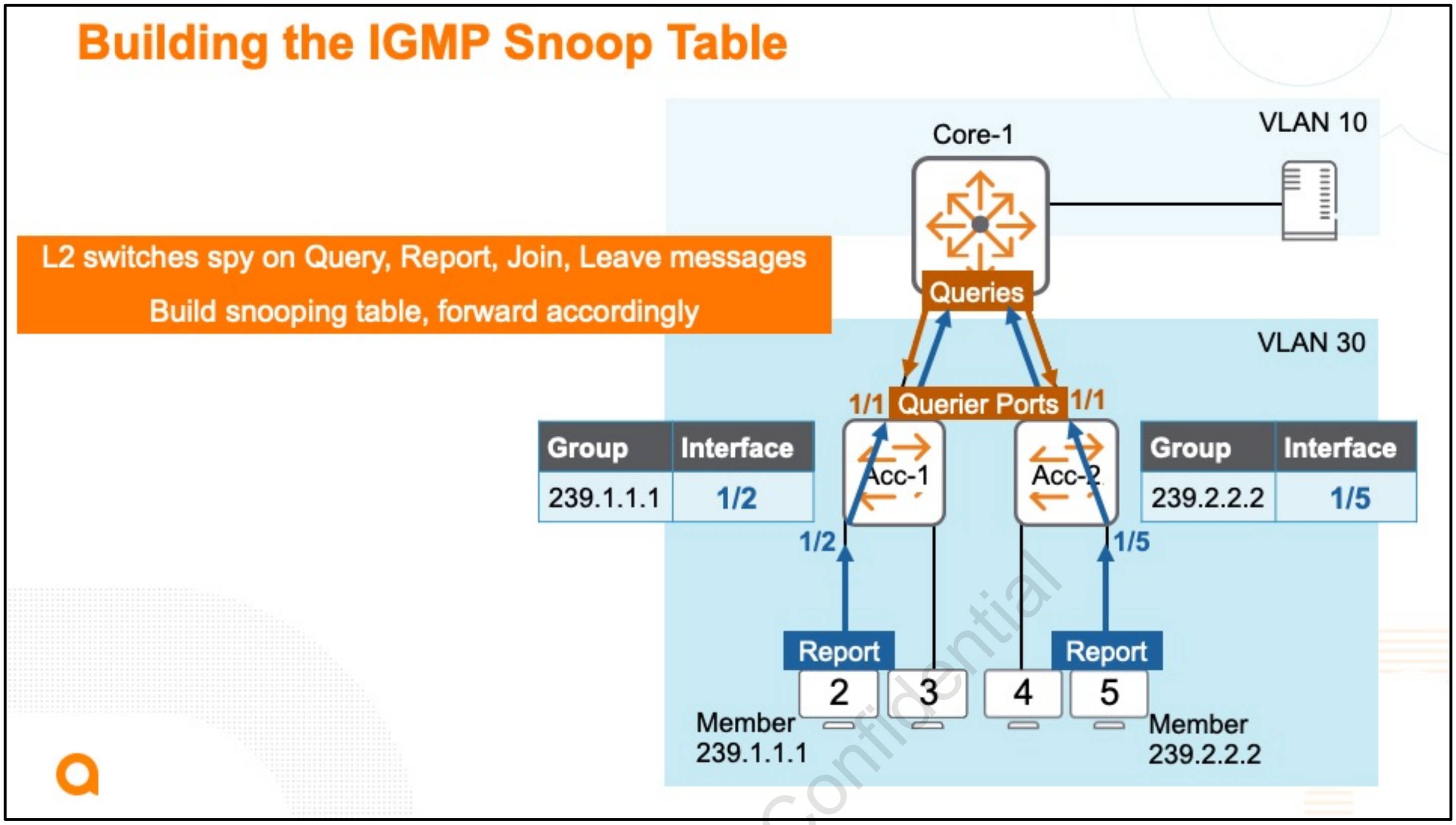
MOD 1- 30

IGMP snooping solves this issue.

| You configure IGMP snooping on Access-1 and Access-2. These Layer 2 devices are transparent to the L2 communications between routers and hosts – they do not run the IGMP protocol, and so act as neither queriers nor hosts. They just quietly listen and learn.

It listens in on (snoops) the normal IGMP packets between querier routers and receiver hosts.

| It learns the MAC addresses of group members and creates Layer 2 multicast forwarding entries - one table for each VLAN interface on which IGMP snooping is enabled. No more traffic to non-member hosts!



Building the IGMP Snoop Table

L2 switches spy on Query, Report, Join, Leave messages

Build snooping table, forward accordingly

Group	Interface
239.1.1.1	1/2

Group	Interface
239.2.2.2	1/5

Here's how switches build a snooping table:

The switch tracks the ports that receive IGMP general queries. These are called querier ports – where it forwards host IGMP reports, joins, and leave messages, and it forwards querier multicast streams. In the figure, port 1/1 on both access switches are Querier ports.

When the switch receives a host IGMP report, it maps the incoming port to the group in the report in the appropriate VLAN IGMP snooping table. Or, if the port is already mapped to that group, the switch resets the dynamic port member aging timer for that port and group. In the figure, port 1/2 maps to 239.1.1.1, and port 1/5 maps to group 239.2.2.2

When hosts hear IGMP reports from other hosts in the same group, they typically suppress their own reports to minimize floods or reports. However, the IGMP snooping switch requires more granular information about the receivers on each port. Therefore, the switch only forwards IGMP reports on router or querier ports. Other hosts, not hearing the report, send their own reports, and the switch can learn their group memberships.

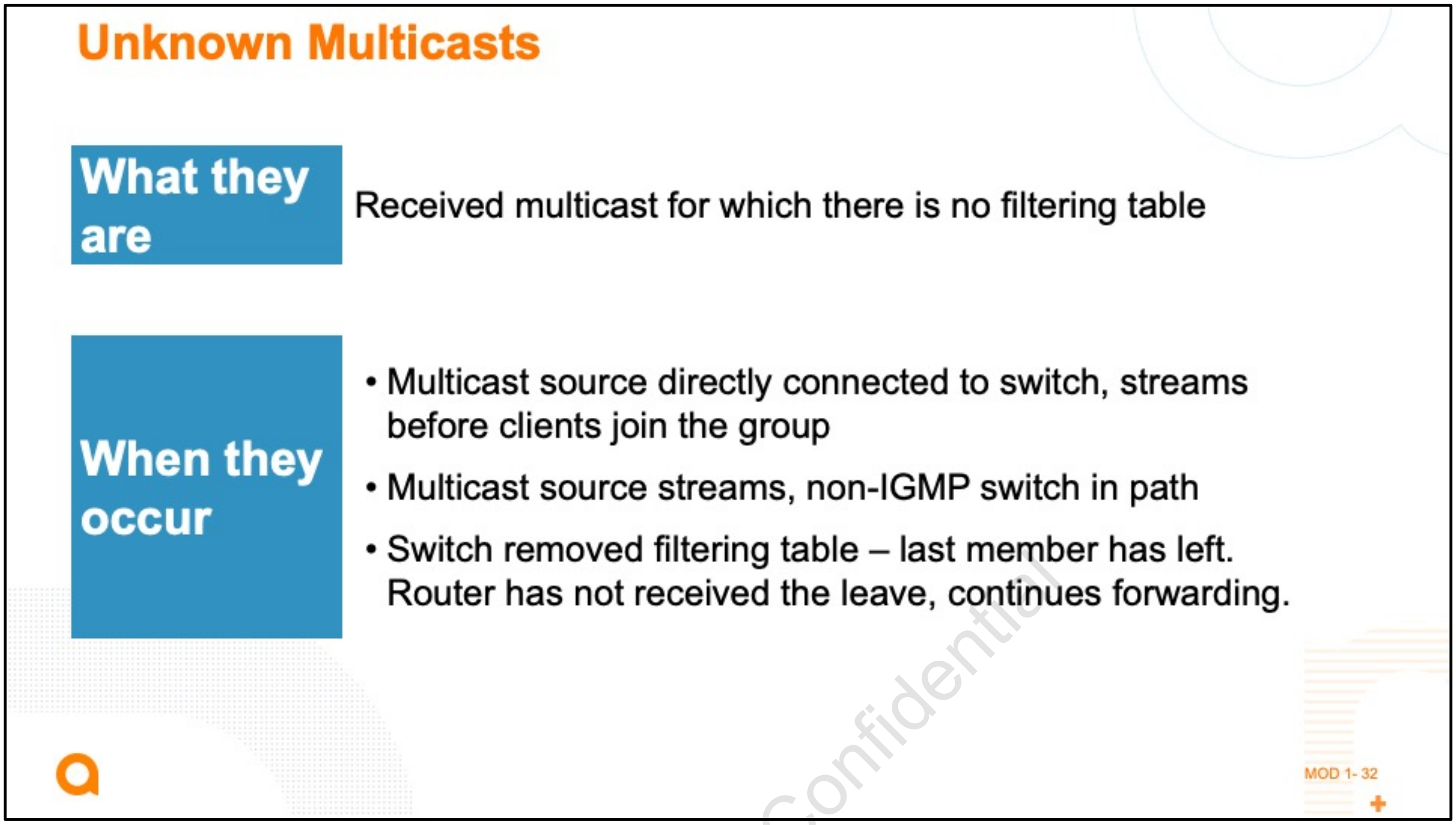
The snooping switch minimizes IGMP report floods – it only forwards the initial entry for a multicast group in its IGMP snooping table (rather than every time it adds a new port to the entry).

Inactive group memberships time out. How the switch handles leave messages depends on the fast-leave setting, which you will examine in a moment.

This is how the IGMP snooping switch builds its table. When the switch receives multicast packets for a particular group on a router port, it forwards them only on the ports mapped to the group.

Note: The switch updates its forwarding table based on the Layer-2 multicast MAC addresses (01:00:5e:xx:xx:xx) and makes forwarding decisions based on what it learned via snooping. Thus, the Layer-2 switches (Access 1 and 2) spy on the query, report, join, and leave messages to learn where the multicast sources and destinations reside and therefore where the multicast stream (239.1.1..1) needs forwarded.

IGMP snooping is disabled by default. The figure shows how to enable IGMP snooping.



Unknown Multicasts

What they are

Received multicast for which there is no filtering table

When they occur

- Multicast source directly connected to switch, streams before clients join the group
- Multicast source streams, non-IGMP switch in path
- Switch removed filtering table – last member has left. Router has not received the leave, continues forwarding.

MOD 1- 32

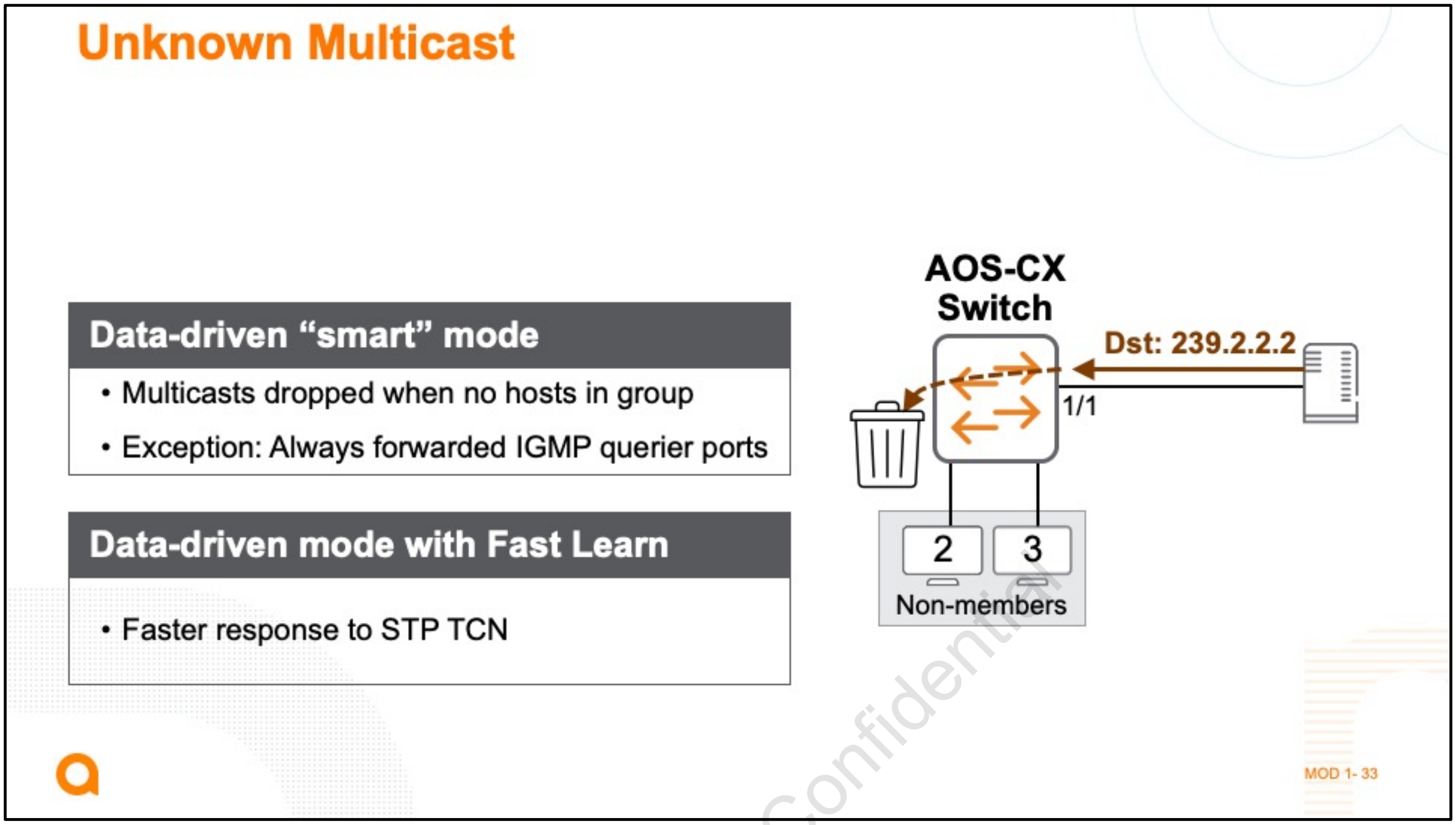
Sometimes an IGMP or IGMP snooping switch receives a multicast for a group for which the switch does not have a filtering table. Called “unknown multicasts,” these multicasts can occur when:

| The multicast source is directly connected to this switch, and it starts streaming before any clients join the group.

A multicast source is streaming, and another switch in the path between the source and this switch does not support IGMP.

The switch has removed a multicast group filtering table because the last member left the group. However, a multicast router forwarding the traffic has not yet received the leave, so it continues to forward the multicasts.

You should understand how this unknown multicast traffic is handled.



Unknown Multicast

Data-driven "smart" mode

- Multicasts dropped when no hosts in group
- Exception: Always forwarded IGMP querier ports

Data-driven mode with Fast Learn

- Faster response to STP TCN

AOS-CX switches operate in data-driven (or smart) mode. They filter all multicasts for which they do not have a multicast table. If no host has joined the group, but the switch receives multicasts, the switch drops them.

One exception applies: an IGMP snooping switch DOES forward both unknown and known multicasts on any port on which it has heard IGMP queries – port 1/1 in the figure. This behavior is required so that a multicast stream can reach the multicast router. In fact, by default, the switch forwards unknown multicasts that arrive on one VLAN on any port on which it has heard queries in any VLAN.

To change this behavior and restrict forwarding to querier ports for that specific VLAN, enter the command:

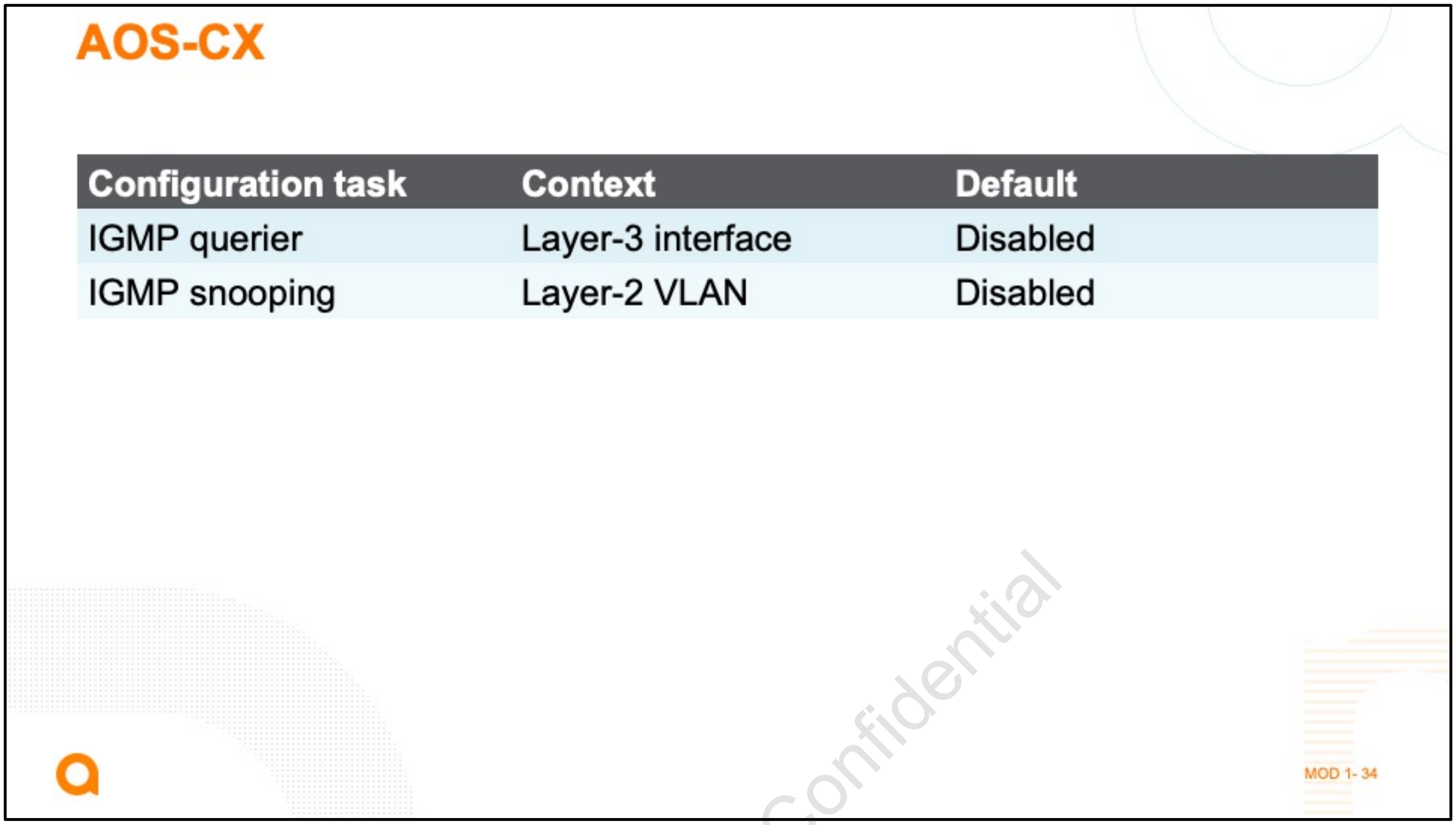
```
Switch(config)# ip igmp snooping drop-unknown vlan-exclusive
```

The default mode is vlan-shared. It is important that you understand that the igmp filter-unknown-mcast command has no effect on how the data-driven switch behaves with non-querier ports. The switch always filters unknown multicasts on these ports.

Data-driven mode and fast learn

Configuring fast learn on a port enables faster response to topology change notifications (TCN). When spanning tree changes the port state from blocked to forwarding, the device acting as querier will immediately send a general query on the fast learn enabled port. Then the device acting as a non-querier will replay the joins. This will help in faster convergence of multicast flows.

```
Switch(config)# vlan <ID>
Switch(config-vlan)# ip igmp snooping fastlearn <port_ID_list>
```



AOS-CX

Configuration task	Context	Default
IGMP querier	Layer-3 interface	Disabled
IGMP snooping	Layer-2 VLAN	Disabled

MOD 1- 34

With the factory default setting, multicast data transmitted from the sources will be flooded on all ports in the VLAN. Configuring IGMP snooping avoids flooding and causes the switch to forward data only to the receivers.

IGMP configuration considerations include the following:

For IGMP to be operational, the interface must be administratively up. For interface VLANs, the L2 VLAN must be up and one of the ports in the VLAN must be up.

The IP address must be assigned for the interface to become querier. Without an IP address, the device will remain in a non querier state.

A querier is required for proper IGMP operation. For this reason, you must enable IGMP on the L3 Interface. If the querier functionality is not configured or disabled, you must ensure that there is an IGMP querier in the same VLAN.

For IGMP snooping to be operational on a VLAN, the VLAN has to be administratively up and at least one port in the VLAN has to be up.

If the switch becomes the querier for a particular interface, then subsequently detects queries transmitted from another device on the same VLAN, the switch ceases to operate as the querier for that interface.

The switch automatically ceases querier operation in an IGMP-enabled interface if it detects another querier on the interface. You can also use the switch CLI to disable the querier capability.

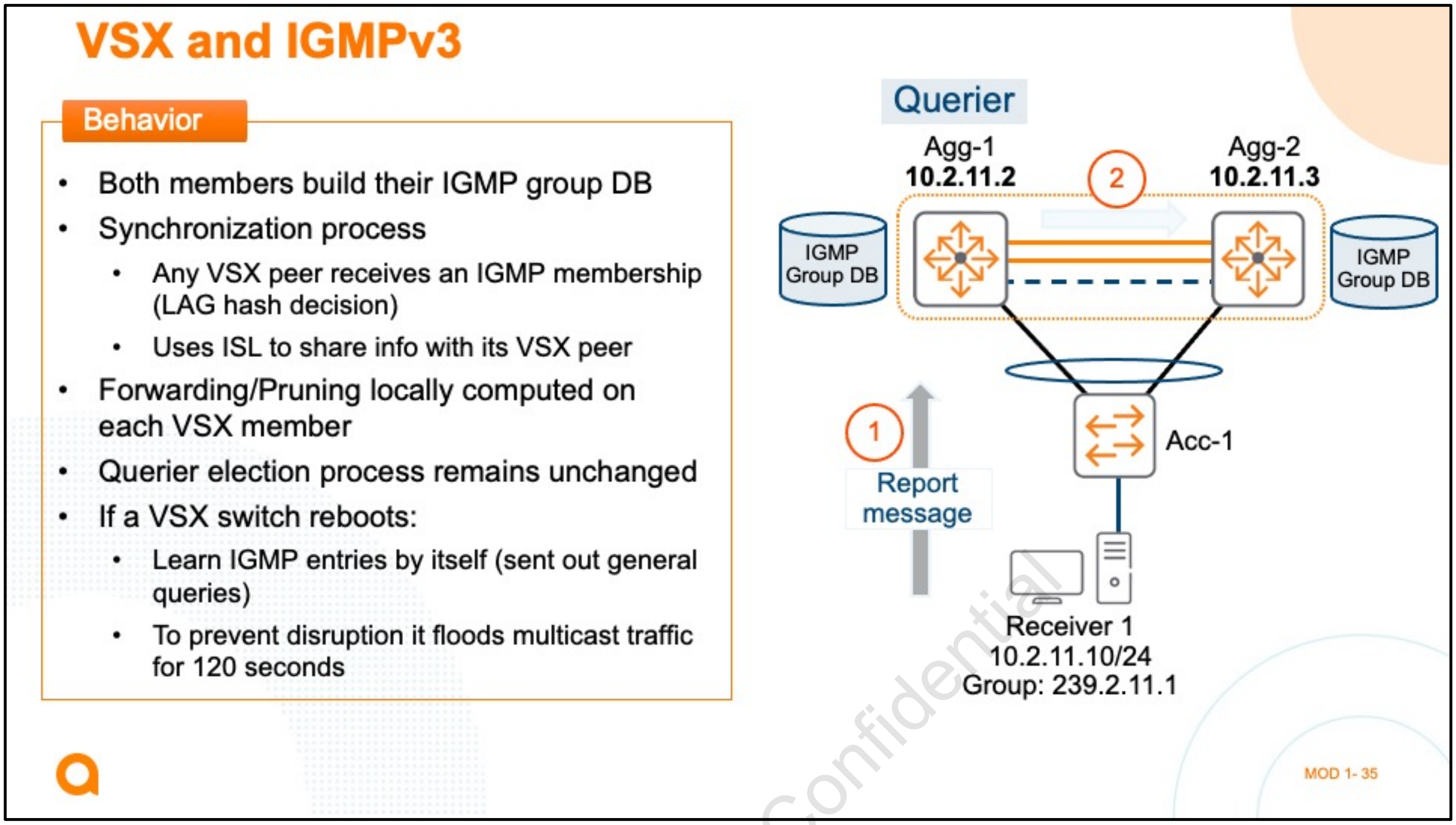
Each of the AOS-CX switches have different capabilities regarding the number of multicast steams supported by the switch. Here’s a summary:

8400 switches: AOS-CX supports 16K IPv4 multicast groups.

8320/8325 switches: AOS-CX supports 4K IPv4 multicast groups.

6400/6300 switches: AOS-CX supports 4K IPv4 multicast groups.

Important: By default, though, switches will treat multicasts as “broadcasts” and flood them in a VLAN unless IGMP snooping is enabled. This wastes bandwidth on segments in a VLAN. Thus, any time multicast is implemented, IGMP snooping should be enabled for VLANs on AOS-CX switches.



VSX and IGMPv3

Behavior

- Both members build their IGMP group DB
- Synchronization process
  - Any VSX peer receives an IGMP membership (LAG hash decision)
  - Uses ISL to share info with its VSX peer
- Forwarding/Pruning locally computed on each VSX member
- Querier election process remains unchanged
- If a VSX switch reboots:
  - Learn IGMP entries by itself (sent out general queries)
  - To prevent disruption it floods multicast traffic for 120 seconds

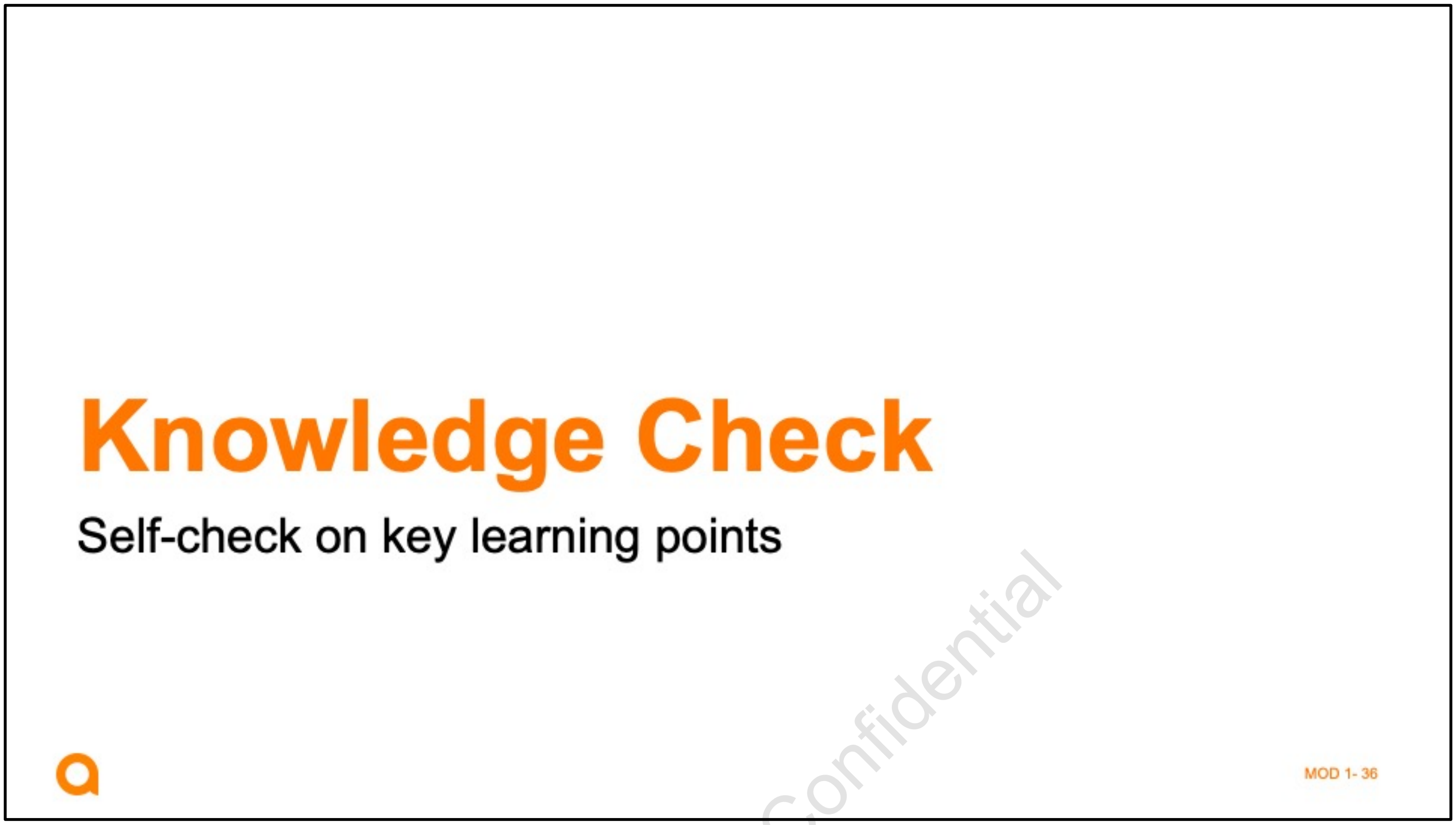
MOD 1- 35

You should also understand how VSX affects the forwarding of multicast traffic. Each VSX switch constructs its own IGMP group database, which is identical to its peer's IGMP group database. Both switches hear JOIN/LEAVE messages they receive from the downstream VSX LAGs because the ISL is always included as a forwarding port for IGMP. For example, SW1 has a VSX LAG to the aggregation switches.

In this example, it chooses the link to Agg-1 for forwarding the IGMP JOIN, and Agg-1 learns the group on the VSX LAG. The IGMP JOIN crosses the ISL to Agg-2. The VSX IGMP process translates the IGMP JOIN learnt on the ISL into an IGMP JOIN message learnt on the VSX LAG, just as if the JOIN had arrived on the link from SW1. Therefore, Agg-2 also learns the group on the VSX LAG to SW1.

In this way, each VSX switch can determine on which interfaces to forward or prune multicast traffic for particular IGMP groups, based on its own IGMP group database. Because the IGMP database construction relies on this dataplane-based process, ISLP does synchronize IGMP groups between VSX peers. If a VSX switch reboots, it needs to re-learn all of the IGMP groups. For about 120 seconds after rebooting, the VSX switch floods multicast traffic within VLANs that have active, forwarding physical ports, which prevents disruption of multicast forwarding during the learning process. The VSX switch also sends an All Hosts Query (AHQ) message. Clients send JOIN messages in response, so the VSX switch can re-learn the groups and re-create the IGMP group database.

You can configure IGMP snooping on access VLANs. It enables VSX switches operating at Layer 2 on the VLAN to similarly construct identical IGMP group databases.

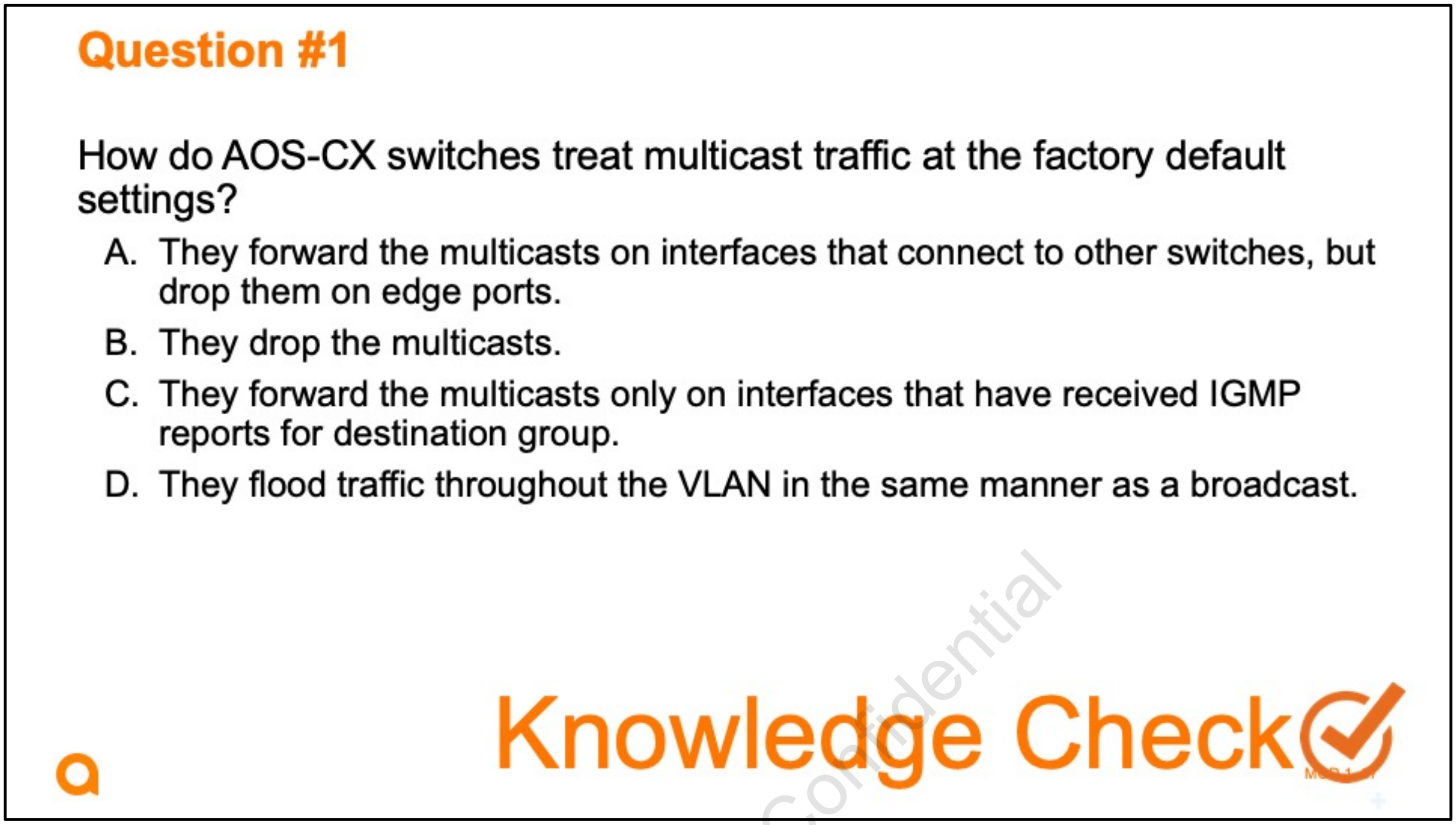


Knowledge Check

Self-check on key learning points

MOD 1- 36

Let’s do a knowledge check.



Question #1

How do AOS-CX switches treat multicast traffic at the factory default settings?

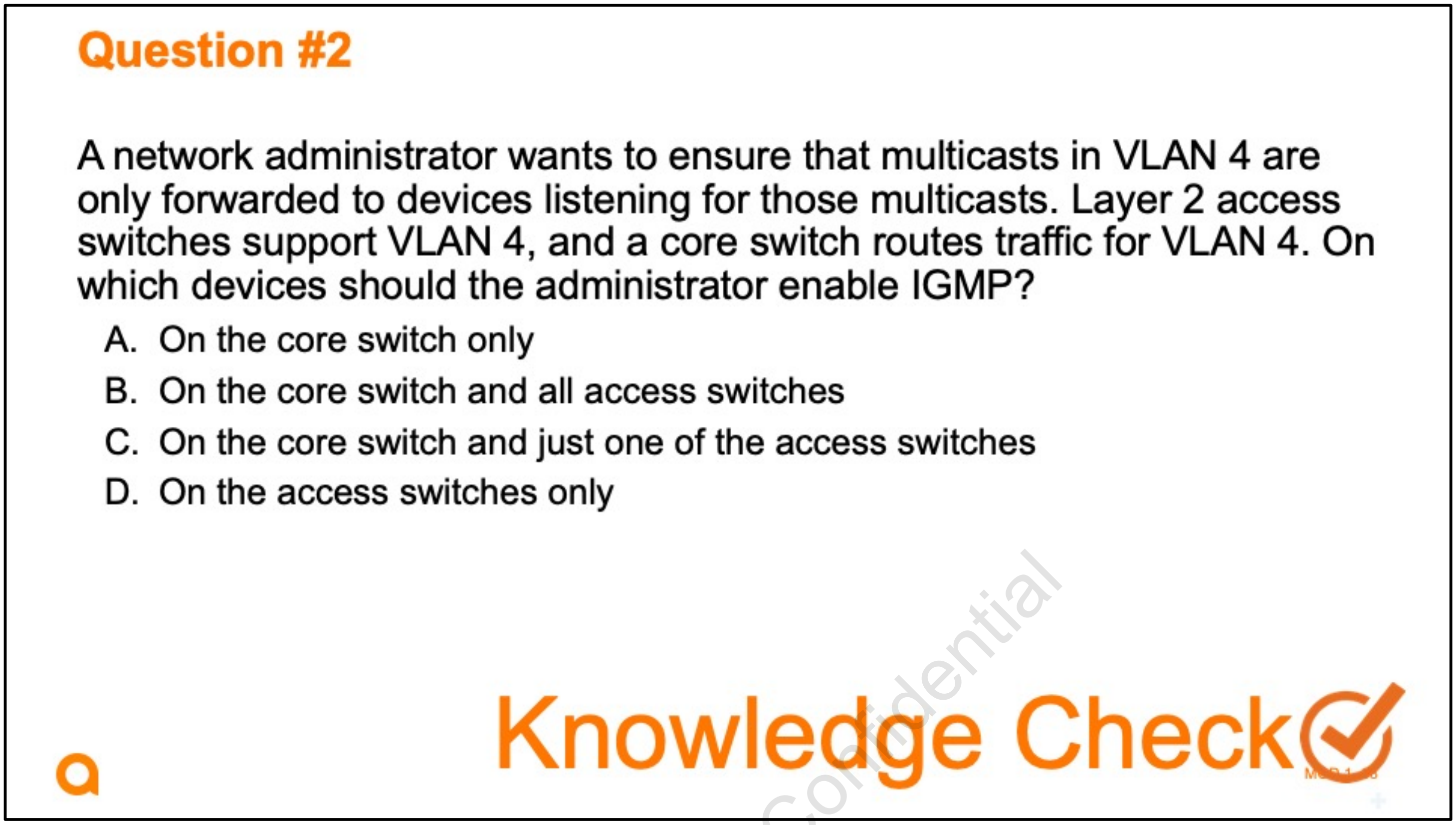
A. They forward the multicasts on interfaces that connect to other switches, but drop them on edge ports.

B. They drop the multicasts.

C. They forward the multicasts only on interfaces that have received IGMP reports for destination group.

D. They flood traffic throughout the VLAN in the same manner as a broadcast.

Knowledge Check



Question #2

A network administrator wants to ensure that multicasts in VLAN 4 are only forwarded to devices listening for those multicasts. Layer 2 access switches support VLAN 4, and a core switch routes traffic for VLAN 4. On which devices should the administrator enable IGMP?

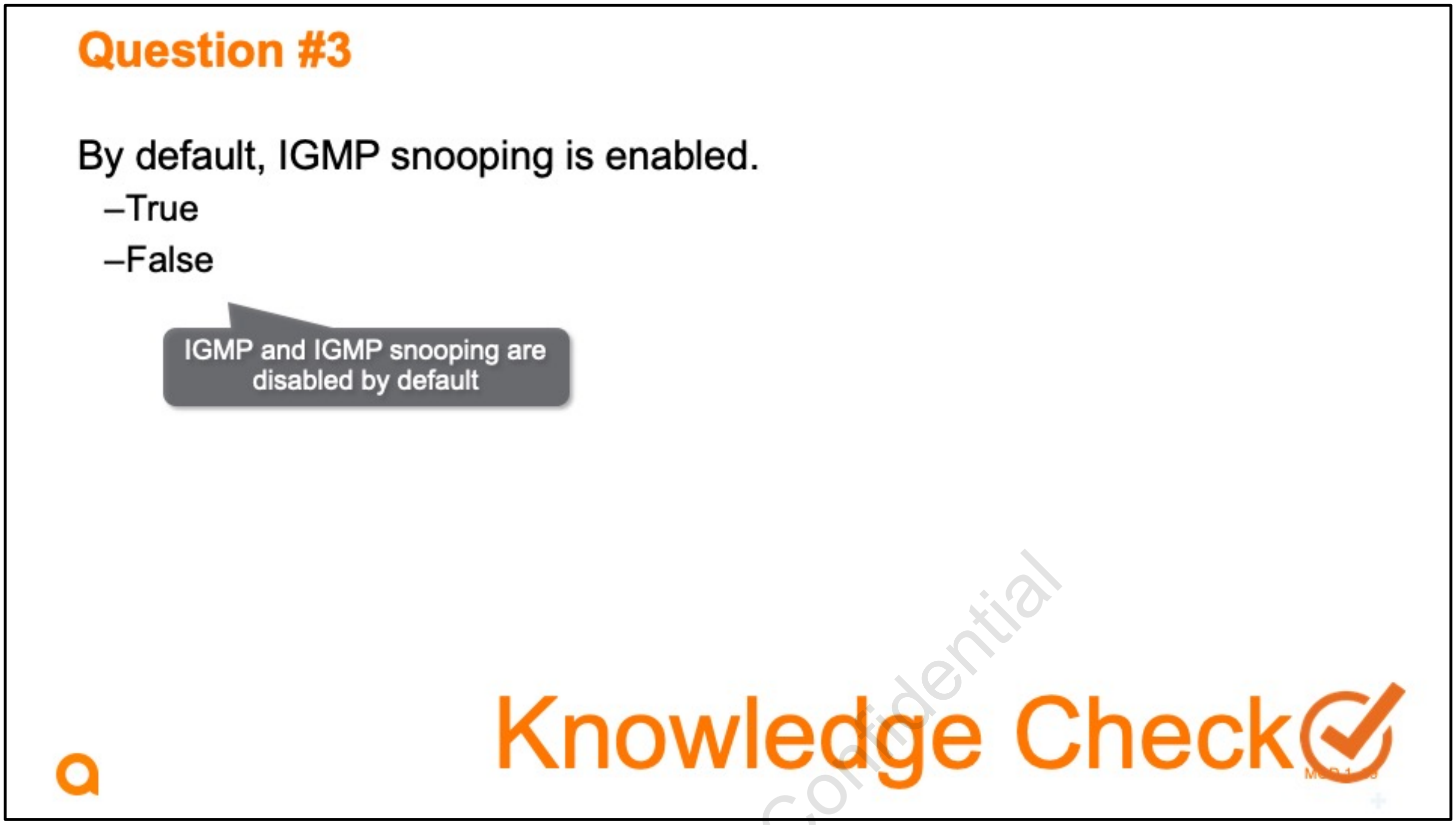
A. On the core switch only

B. On the core switch and all access switches

C. On the core switch and just one of the access switches

D. On the access switches only

Knowledge Check



Question #3

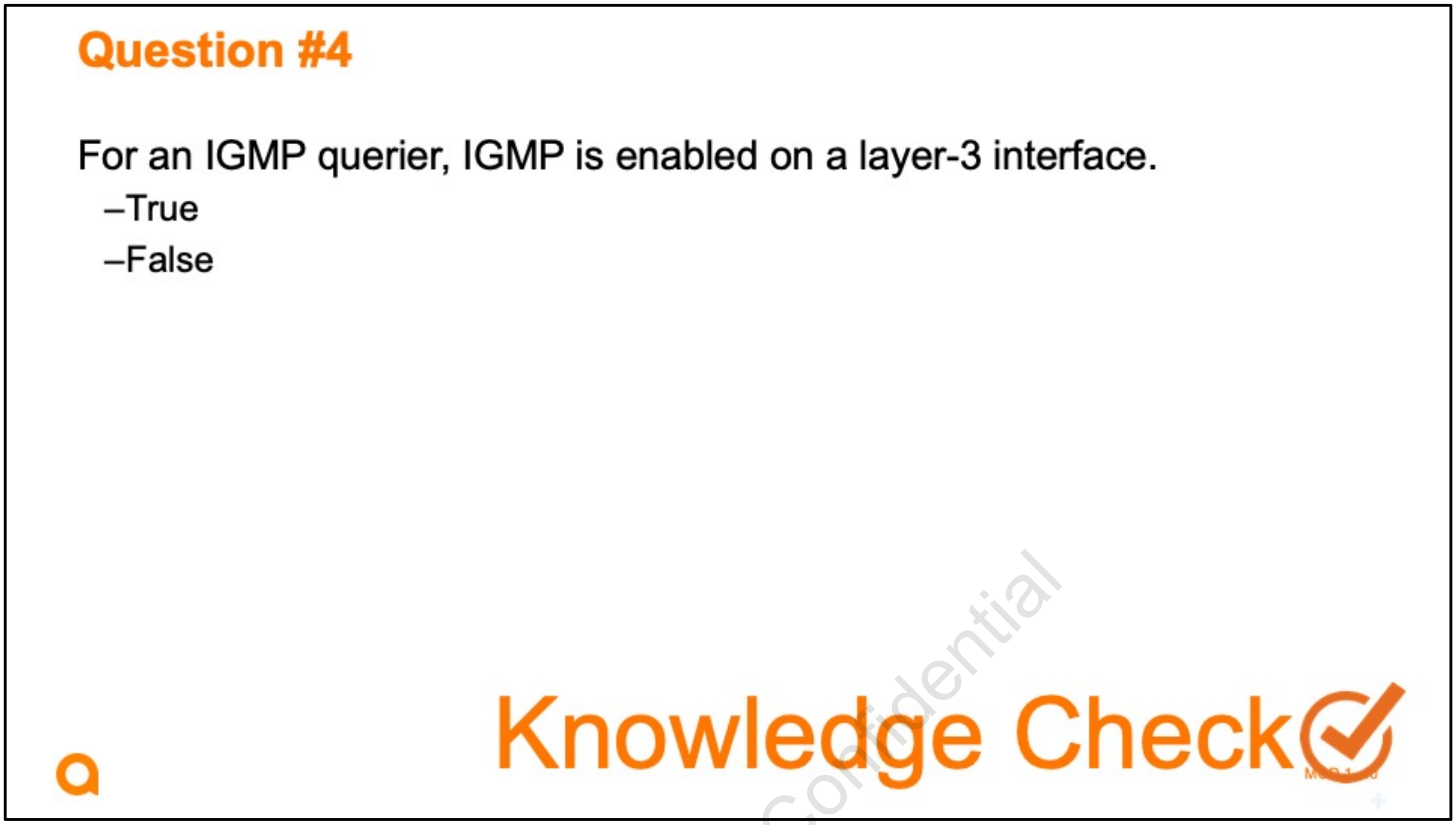
By default, IGMP snooping is enabled.

–True

–False

IGMP and IGMP snooping are disabled by default

Knowledge Check



Question #4

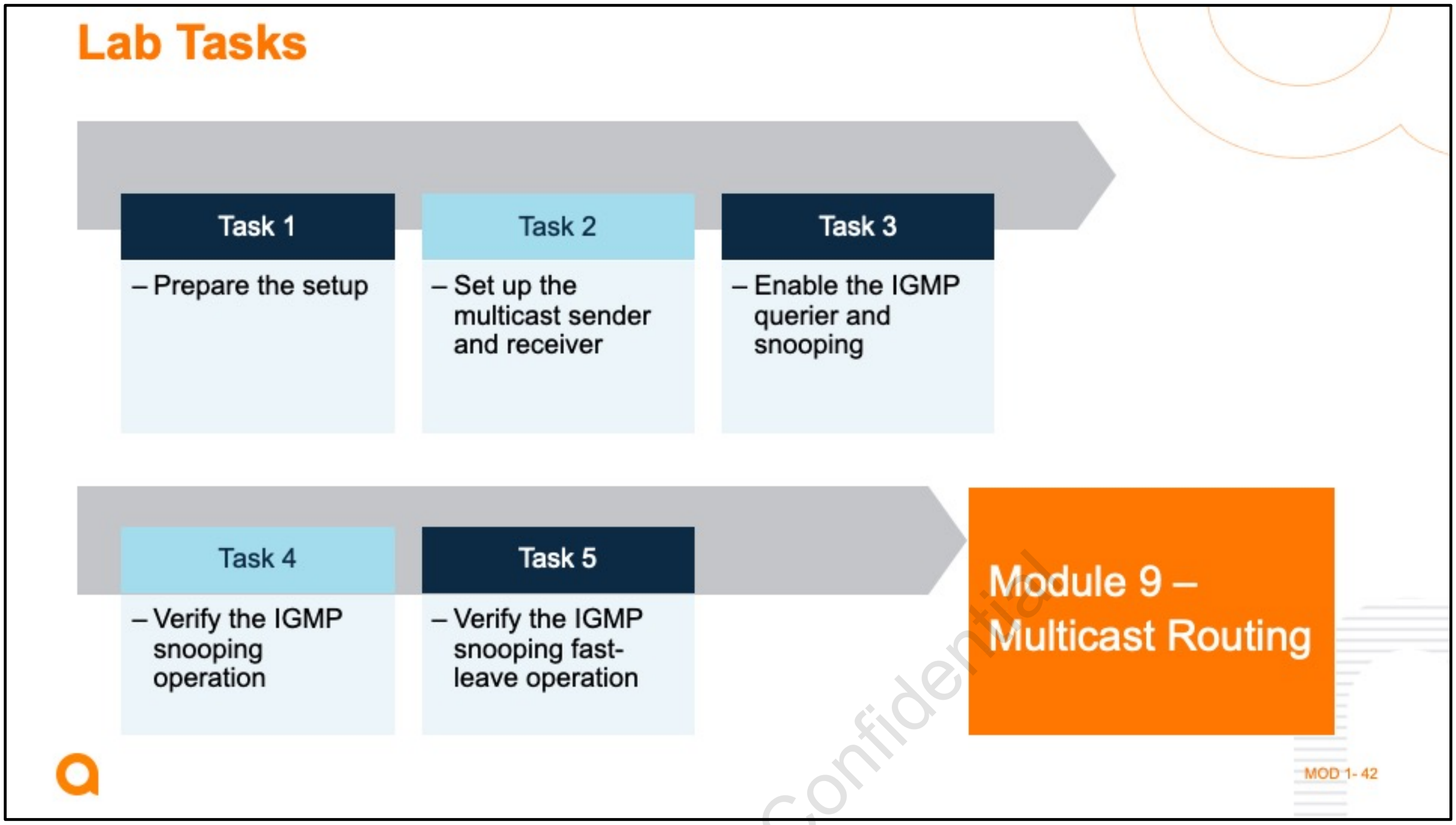
For an IGMP querier, IGMP is enabled on a layer-3 interface.

- –True
- –False

Knowledge Check

Lab Activity

Its time for a lab activity.



Lab Tasks

Task 1
– Prepare the setup

Task 2
– Set up the multicast sender and receiver

Task 3
– Enable the IGMP querier and snooping

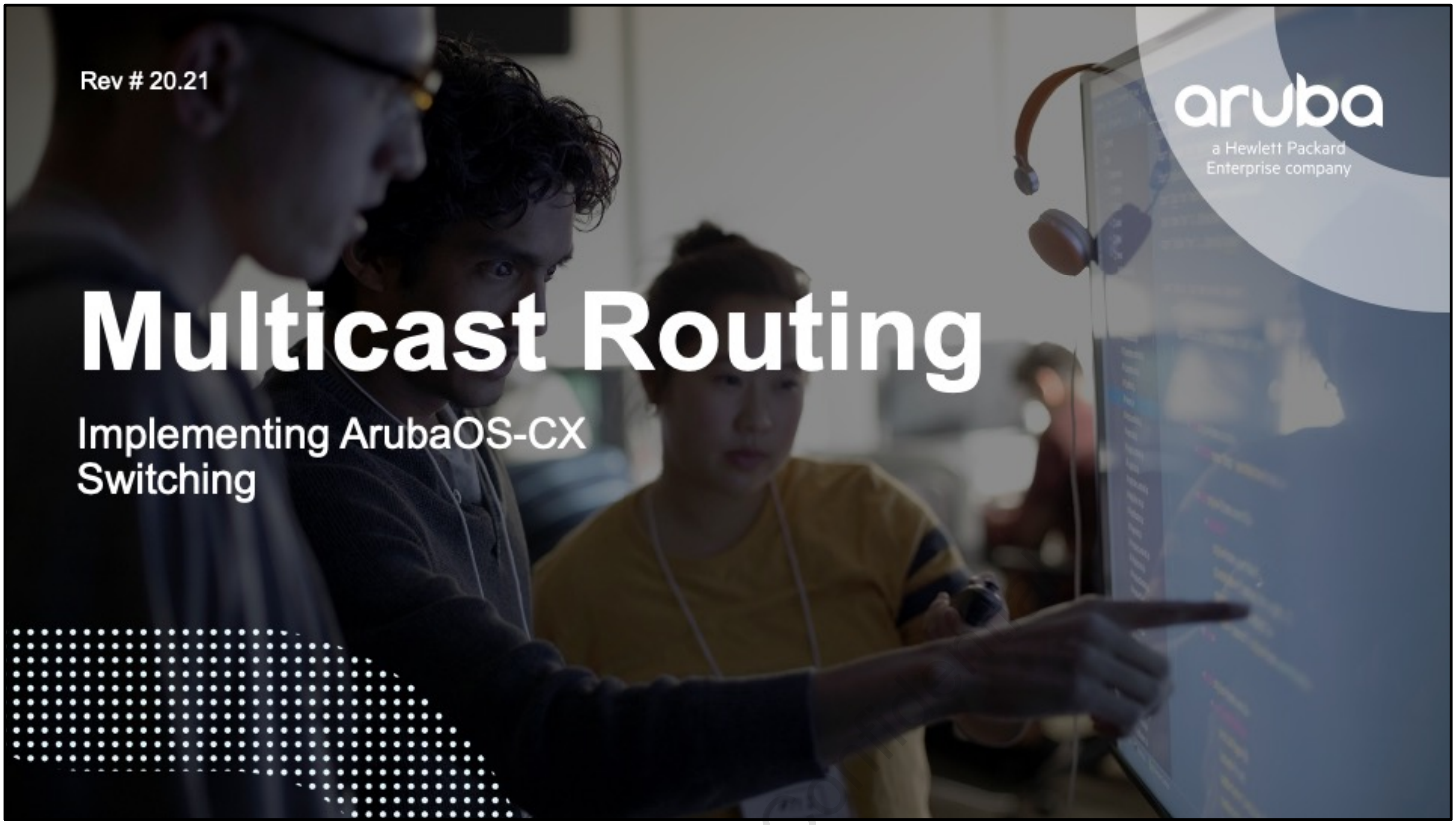
Task 4
– Verify the IGMP snooping operation

Task 5
– Verify the IGMP snooping fast-leave operation

Module 9 – Multicast Routing

MOD 1- 42

The figure provides a brief review of lab tasks. Please see your lab guide for details. When you are ready, please continue with Module 9 – Multicast Routing.

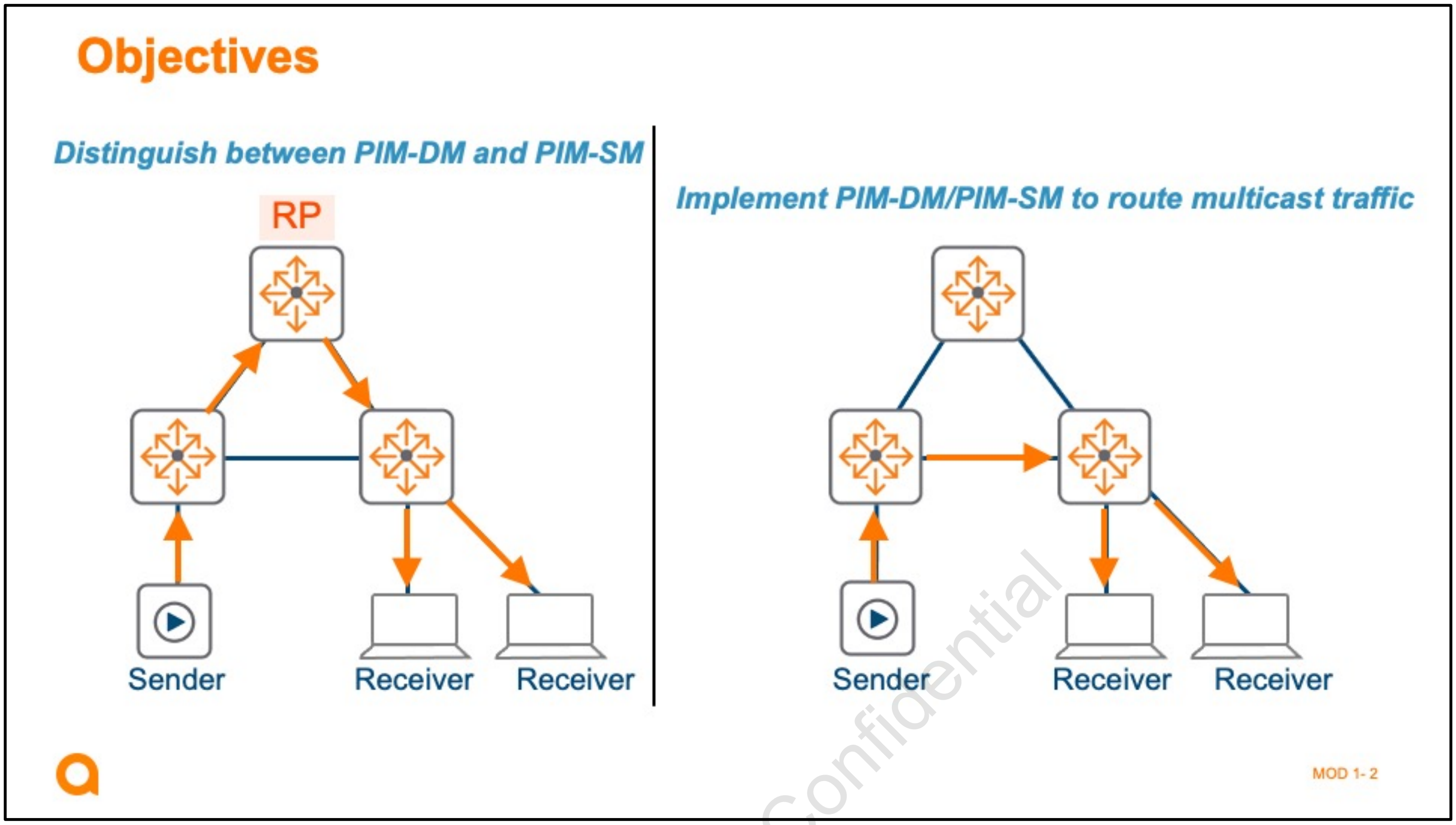


Rev # 20.21

Multicast Routing

Implementing ArubaOS-CX Switching

You learned IGMP in the last Module, so we into Multicast Routing.



Objectives

Distinguish between PIM-DM and PIM-SM

Implement PIM-DM/PIM-SM to route multicast traffic

MOD 1- 2

After completing this module, you will be able to:

Distinguish between PIM-DM and PIM-SM

Implement PIM-DM and PIM-SM to route multicast traffic

Overview

PIM Intro

PIM-DM

PIM-SM

Lab Activity

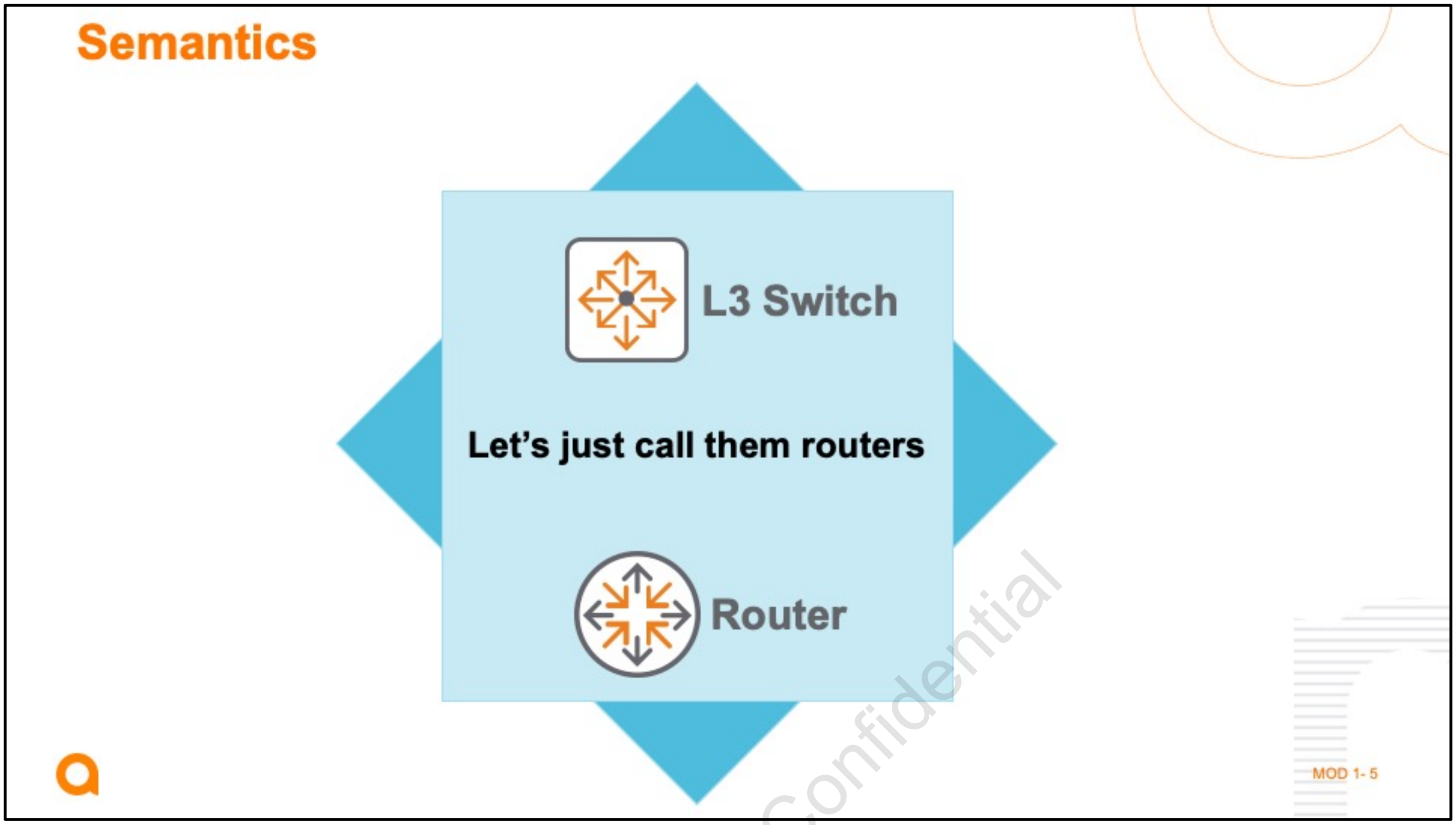
MOD 1- 3

In this module, you will be introduced to Protocol Independent Multicast (PIM), a multicast routing protocol that enables switches to determine which interfaces should receive multicast traffic in a layer-3 network, like a campus.

In the last module, you learned about how AOS-CX switches can learn where the multicast source (typically a server) and the multicast receivers (typically user devices) reside. The problem in a larger campus network is that the multicast source and destinations are commonly on different subnets, which requires a routing function. Unicast routing protocols like OSPF will not route multicast traffic. You need a multicast routing protocol, and PIM is the de facto standard. This module introduces the PIM routing protocol and the two modes of PIM: dense mode and sparse mode.

PIM Introduction

a Hewlett Packard Enterprise company



Semantics

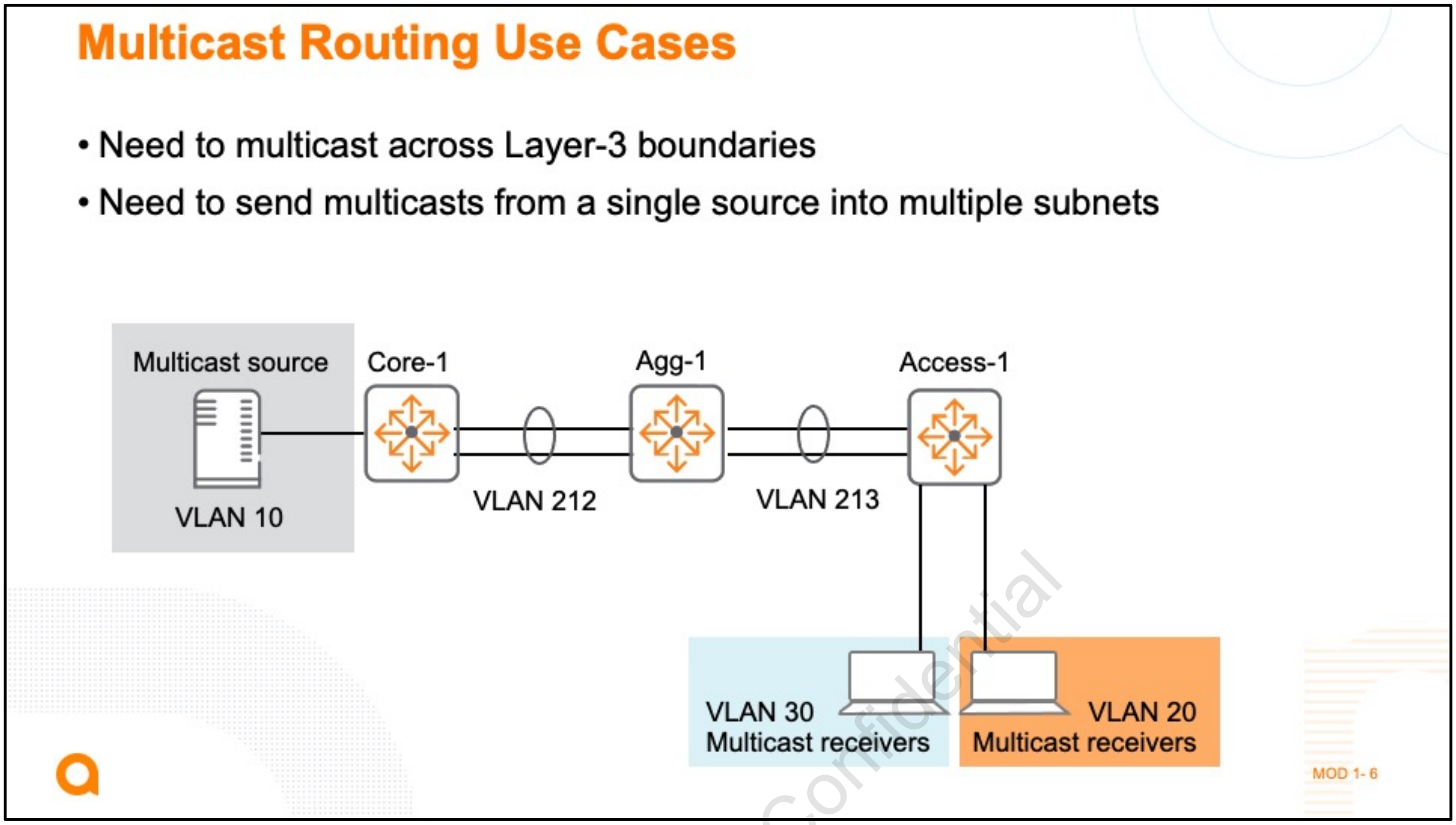
L3 Switch

Let's just call them routers

Router

MOD 1- 5

You know the deal by now - Layer 3 switch…purpose-built router – its all about the same when they're pushing packets, and "router" is only 2 syllables, so I'll just say that.



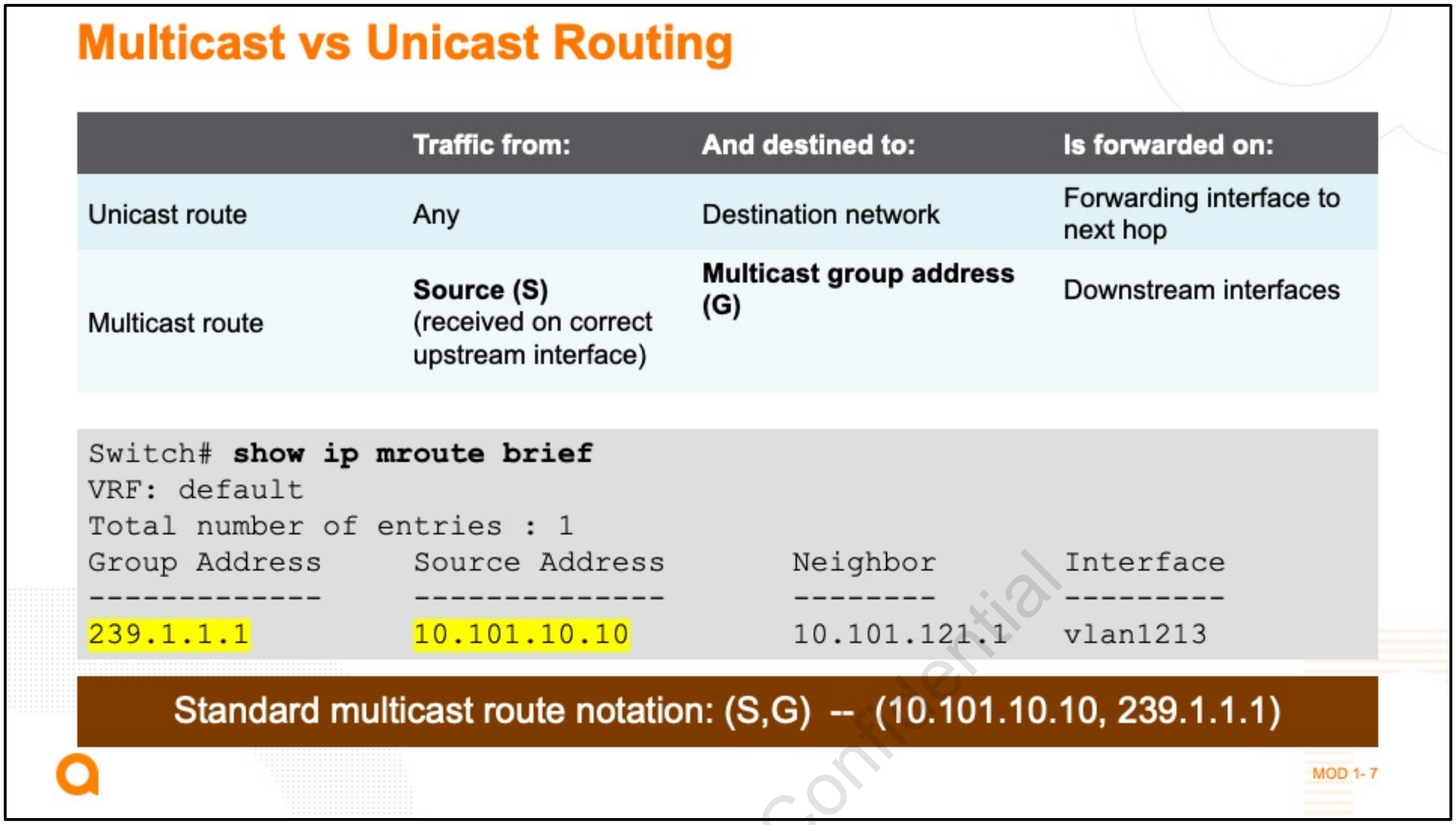
Multicast Routing Use Cases

- Need to multicast across Layer-3 boundaries
- Need to send multicasts from a single source into multiple subnets

For small networks with only 1 router, or a single VLAN, there is no need for a multicast routing protocol like PIM. You just studied scenarios like this in the previous section about IGMP.

However, if you have a more complex, enterprise-class network with multiple routers, you need PIM.

In the figure, the multicast source is in VLAN 10. This traffic must traverse two routers to reach the receivers in VLANs 20 and 30. Multicast routing enables a source in one VLAN to stream to receivers in multiple VLANs.



Multicast vs Unicast Routing

	Traffic from:	And destined to:	Is forwarded on:
Unicast route	Any	Destination network	Forwarding interface to next hop
Multicast route	**Source (S)** (received on correct upstream interface)	**Multicast group address (G)**	Downstream interfaces

```
Switch# show ip mroute brief
VRF: default
Total number of entries : 1
Group Address       Source Address         Neighbor       Interface
-------------       --------------         --------       ---------
239.1.1.1           10.101.10.10           10.101.121.1   vlan1213
```

Standard multicast route notation: (S,G) -- (10.101.10.10, 239.1.1.1)

MOD 1- 7

The figure summarizes the difference between unicast and multicast routing. While unicast routing forwards any traffic toward some destination, multicast traffic forwards traffic away from a specific source, with a multicast group address as its destination.

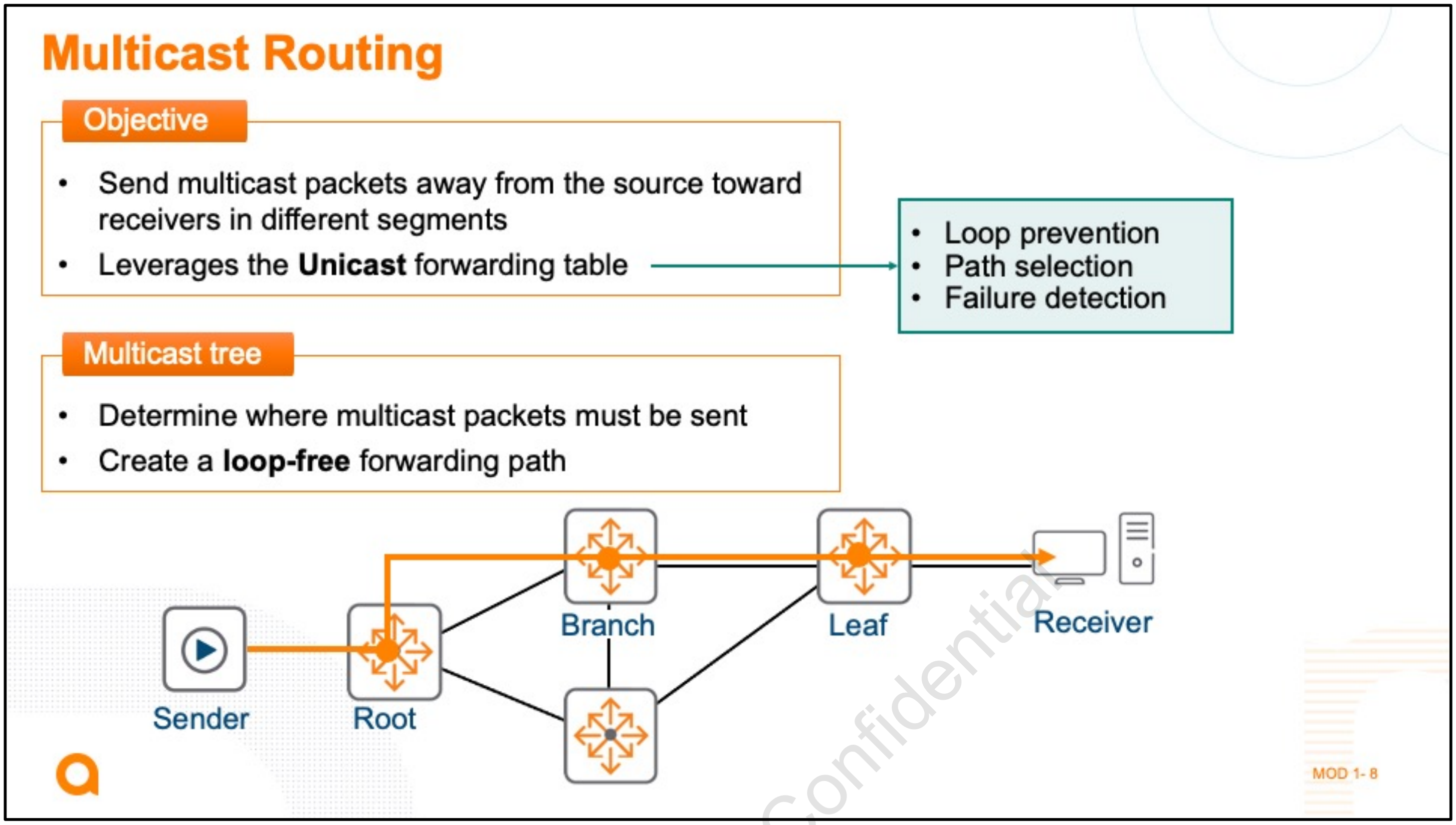
Unicast routing protocols like OSPF do not care about the source address of packets – they compare the destination IP address of each packet to their route table, and forward packets along the best path.

For multicast routing, the source matters. A multicast route must track the upstream interface that is closest to the source so that the router can verify that the multicasts are arriving on the correct interface. It can then drop multicasts that arrive on the wrong interface, preventing multicasts from looping through the network.

The figure shows the output of show ip mroute brief – the multicast route (mroute) table. The router has learned that it should forward all packets for group 239.1.1.1 efficiently away from the source of this stream – 10.101.10.10, toward its neighbor at 10.101.121.1.

An active multicast route is called an (S, G) route because it includes both the source IP address (S) and the destination multicast group address (G). Multicast routing devices create an (S, G) route after they receive a multicast from a source to a group. For example, the (10.101.10.10, 239.1.1.1) route refers to traffic from 10.101.10.10 to 239.1.1.1.

You will soon learn about (*,G) entries, which act as a kind of placeholder until the multicast source is discovered.



Multicast Routing

Objective

- Send multicast packets away from the source toward receivers in different segments
- Leverages the **Unicast** forwarding table

- Loop prevention
- Path selection
- Failure detection

Multicast tree

- Determine where multicast packets must be sent
- Create a **loop-free** forwarding path

You know that regular unicast routing protocols like OSPF send packets toward a destination. The focus is on each packet's destination IP address. The objective of multicast routing is to send multicast messages away from the source toward all interested receivers. Multicast packets must always flow away from the source, and never back on a segment from where the transmission originated. This means that rather than tracking only destinations, multicast routers must also track the location of sources, the inverse of unicast routing. Even though multicast uses the exact inverse logic of unicast routing protocols, you can leverage the information obtained by those protocols (Unicast routing – OSPF and BGP for example) for multicast forwarding.

This allows a network to avoid using another full routing protocol, minimizing the use of memory space. All the functionality that is built into the unicast routing protocol, including loop prevention, path selection, failure detection and so on, can be used for multicast. Modern IP multicast routing uses Protocol Independent Multicast (PIM), which leverages unicast routing information, to forward messages to receivers. The process of determining where the multicast messages are to be sent is referred to as building the tree.

With IPv4, IGMP determines who needs to receive a multicast stream, PIM helps determine how to get the stream to those destinations. relies on IGMP to request the traffic. IPv6 uses MLD instead of IGMP.

Note: If you want to route multicast traffic into a VLAN, it is recommended that you avoid multinetting on that VLAN (configuring more than one subnet).

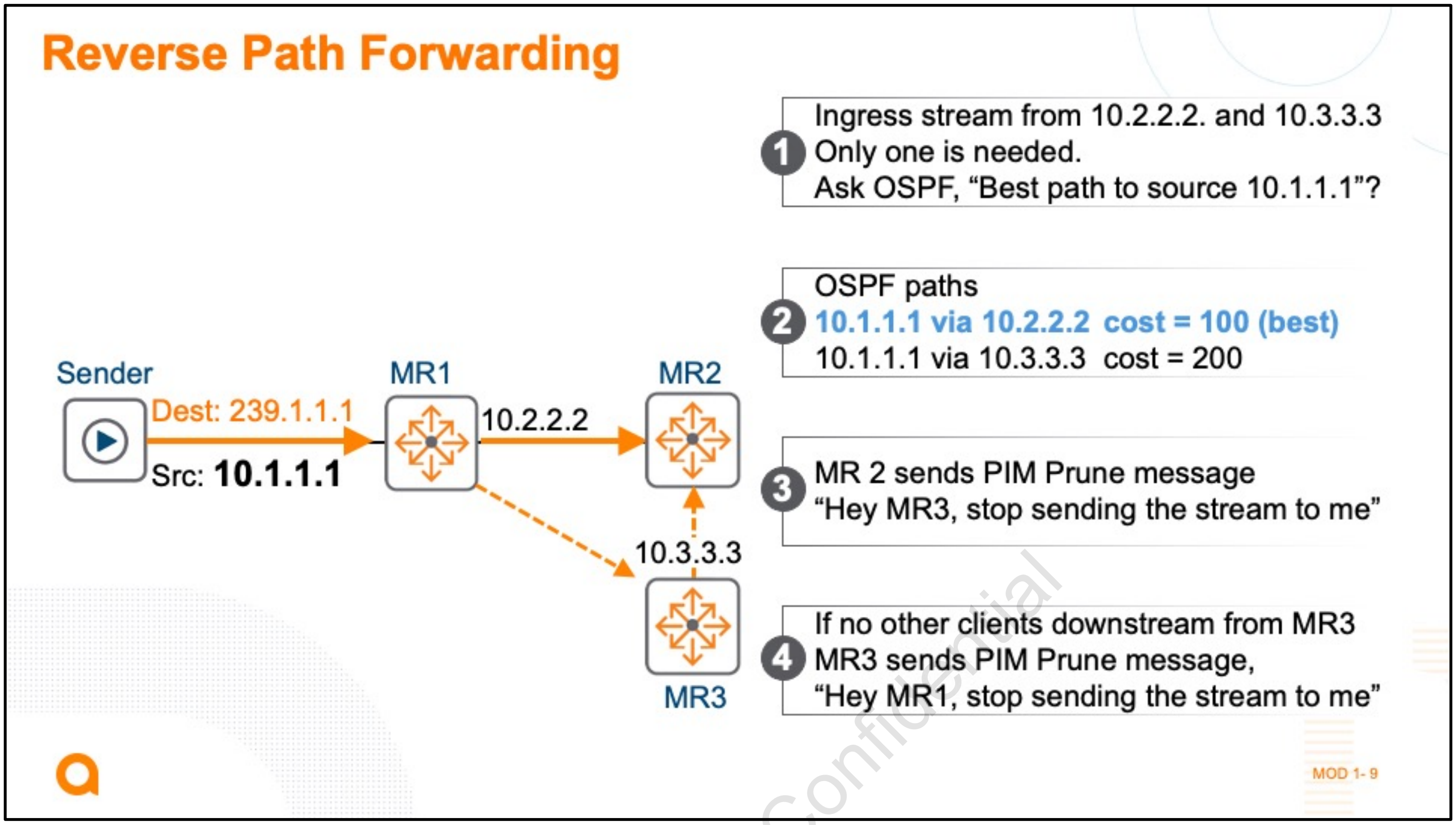
Multicast tree

A network tree is essentially a graph, or directed graph (digraph), that represents a hierarchical logical structure. The primary purpose of a multicast tree is to build an efficient loop-free forwarding path from the source of the multicast stream toward all the receivers. A multicast tree is composed by the following components:

Root. Router closest to the source.

Leaf. Router with an attach receiver.

Branch. Intermediate router that performs replication to connect the root to the leaves.



Reverse Path Forwarding

Sender

Dest: 239.1.1.1

Src: **10.1.1.1**

MR1

10.2.2.2

MR2

10.3.3.3

MR3

1 Ingress stream from 10.2.2.2. and 10.3.3.3
Only one is needed.
Ask OSPF, "Best path to source 10.1.1.1"?

2 OSPF paths
10.1.1.1 via 10.2.2.2 cost = 100 (best)
10.1.1.1 via 10.3.3.3 cost = 200

3 MR 2 sends PIM Prune message
"Hey MR3, stop sending the stream to me"

4 If no other clients downstream from MR3
MR3 sends PIM Prune message,
"Hey MR1, stop sending the stream to me"

MOD 1- 9

PIM uses a so-called Reverse Path Forwarding (RFP) technique to save bandwidth and maximize efficiency. Here's the scenario. A multicast source originates a stream with source address 10.1.1.1, and multicast destination 239.1.1.1. Multicast Router MR1 receives this stream and forwards it to its two neighbors –MR2 and MR3. MR3 receives this stream and forwards it to its only other neighbor – MR2.

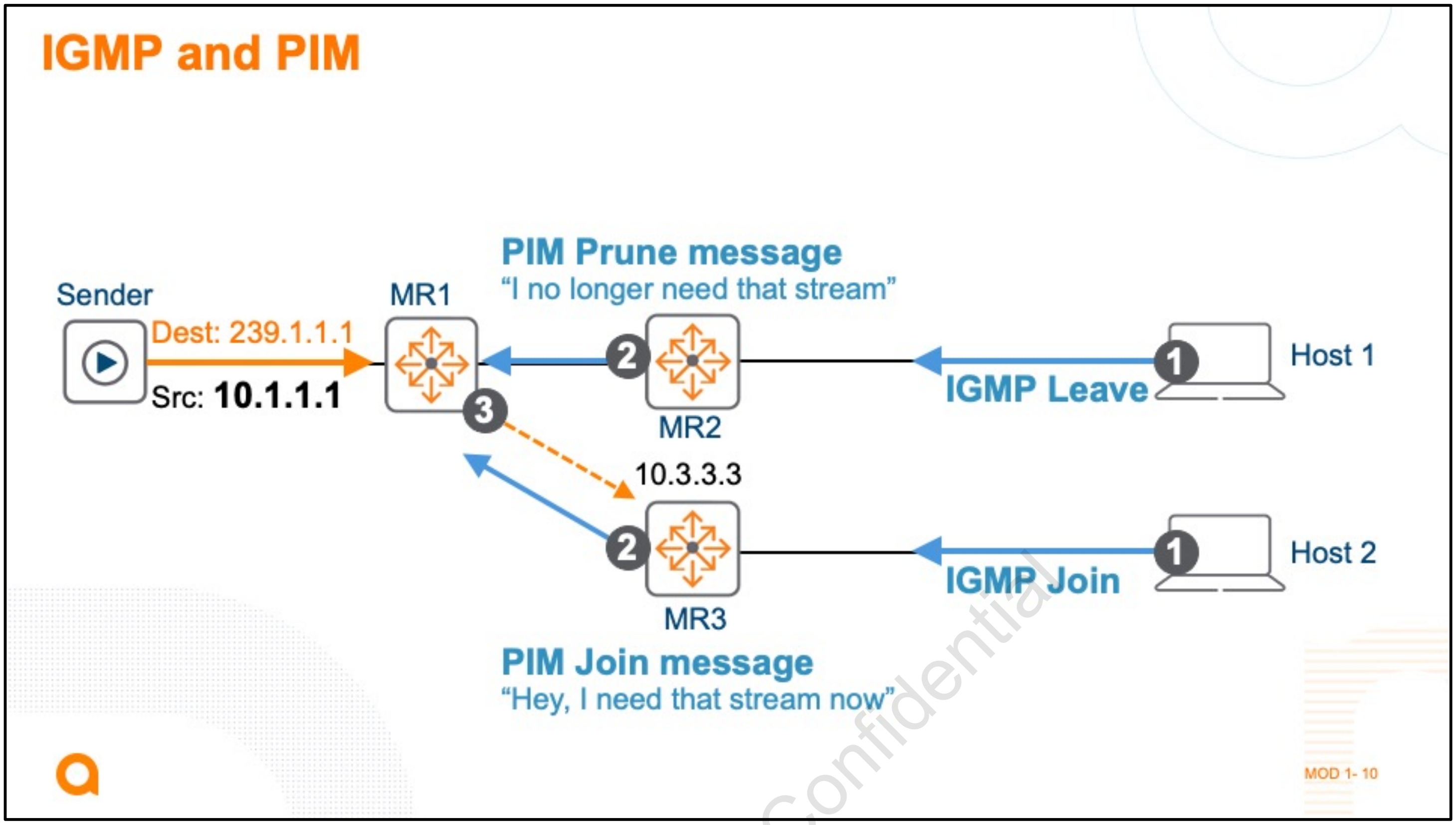
Thus, MR2 receives two copies of the multicast stream. This is a waste of bandwidth! Remember, the advantage of multicasting is that a single stream can reach multiple destinations. One of these streams must be eliminated, So the router looks in the unicast routing table, "What is the best path back to source 10.1.1.1?

MR2's OSPF process has learned that it can reach 10.1.1.1 via 10.2.2.2 (MR1) with a cost of 100, and can reach it via 10.3.3.3 (MR3) with a cost of 200. Thus, 10.2.2.2 is the best path back to 10.1.1.1. Of course, if the best "reverse path" back to 10.1.1.1 is via 10.2.2.2, then that is the best path for receiving multicast traffic from 10.1.1.1.

And so MR2 sends a PIM message to MR3, "Hey MR3, I don't need you to send that stream to me. Please stop". MR3 complies to this request. In other words, the 10.3.3.3 interface is "pruned from the tree". Bandwidth is no longer wasted sending the stream over that interface. If MR3 has no receivers directly attached to it (as determined by IGMP), and host no other downstream routers with receivers, than MR3 has no reason to receive the stream. In that case, MR3 sends a message to MR1, "Please stop sending the stream to me".

MR1 complies and so prunes that interface from the multicast distribution tree. In a large, complex network with dozens or hundreds of routers, every router uses this RFP process. The result is that multicast traffic is only sent where needed. Nice! And, like spanning-tree, if a link failure occurs the PIM routers quickly discover the issue and determine an alternative, loop-free path.

We have an efficient multicast tree for the current situation, but networks are dynamic. New users join a multicast group, existing users leave. How is this handled?



IGMP and PIM

Sender

Dest: 239.1.1.1

Src: **10.1.1.1**

MR1

PIM Prune message

“I no longer need that stream”

2

MR2

1

IGMP Leave

Host 1

3

10.3.3.3

2

MR3

1

IGMP Join

Host 2

PIM Join message

“Hey, I need that stream now”

MOD 1- 10

In the previous example, we had an efficient multicast forwarding tree that only forwarded the stream out needed interfaces. Then things change.

Host 1 is the last remaining receiver downstream of Multicast Router MR2, and it sends an IGMP Leave message, “I no longer need the stream”.

In response, MR2 sends a PIM message to MR1, “I no longer need the stream”. MR1 sees that there are no more downstream neighbors out that interface, and so prunes that interface from the multicast distribution tree.

Meanwhile, some employee sits down at her desk and runs some application that requires the stream. The host sends an IGMP Join message to MR3, “I need that stream”. In response, MR3 sends a PIM message to MR1, “I need that stream”. MR1 add the MR3-connected interface back to the distribution tree, and the stream again flows to MR3, and on to Host 2.

This confirms what you learned earlier:

IGMP is Host-to-router communications, so routers can discover which receivers need a particular stream, and which receivers do not.

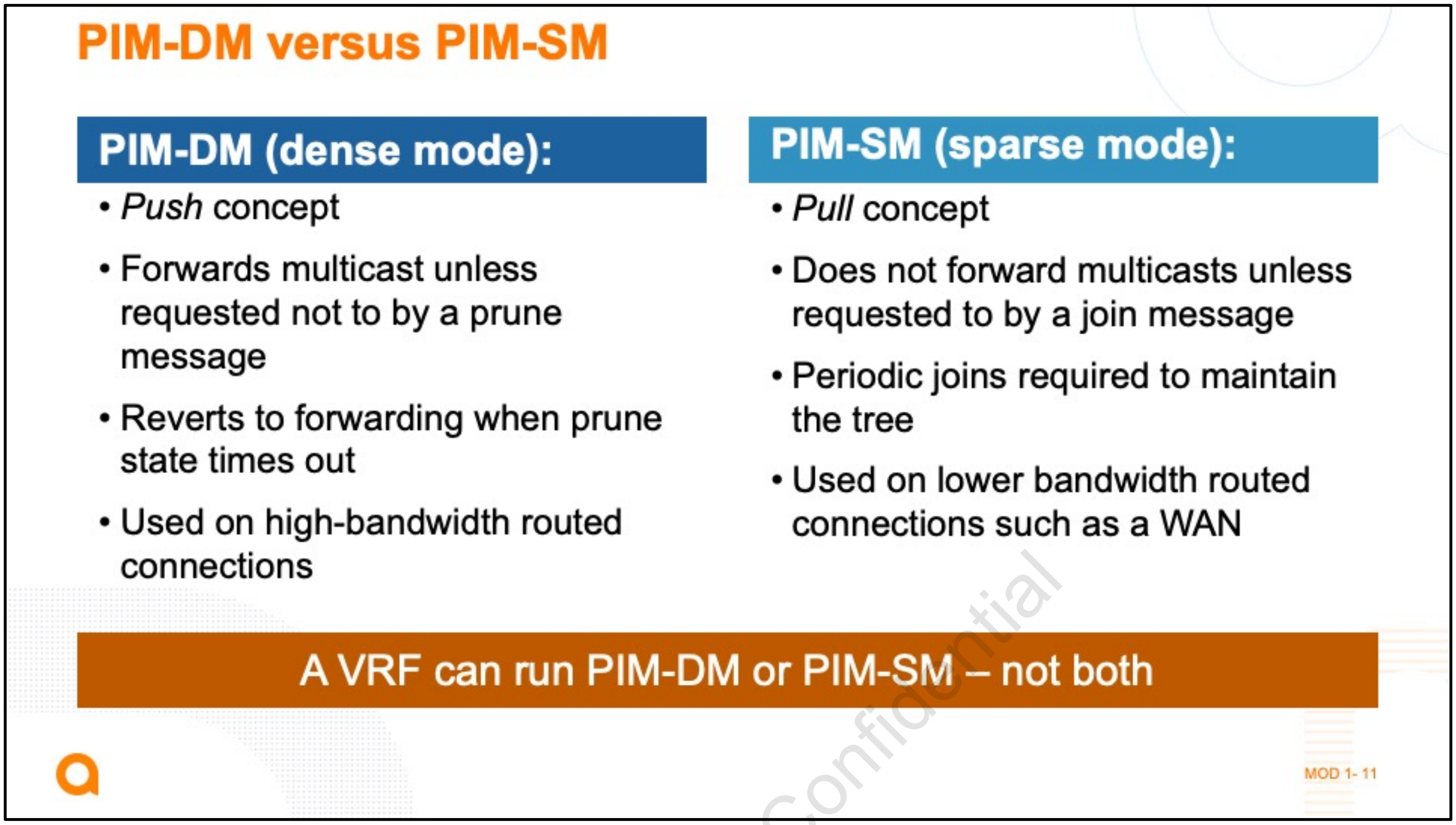
PIM is router-to-router communications, so multiple routers can build a shortest path tree

(SPT) for a multicast stream, leveraging the internal routing protocol (OSPF, etc) for Reverse Path Forwarding.

It is important that you understand that IGMP alone is never enough to provide multicast routing even if the multicast source and multicast receivers are connected to VLANs on the same multicast routing switch. This switch must still have multicast routing enabled and run PIM on the VLAN interfaces to route multicast traffic from one subnet to another.

In other words, IGMP and PIM are a team that enable efficient multicast routing. They build and maintain a dynamic multicast distribution tree. Each device maintains a list of interfaces that are part of the tree, out which a multicast stream is forwarded. Other interfaces are pruned from the tree – removed from this Outbound Interface List (OIL).

There are two methods used to build this distribution tree.



PIM-DM versus PIM-SM

PIM-DM (dense mode):

- *Push* concept
- Forwards multicast unless requested not to by a prune message
- Reverts to forwarding when prune state times out
- Used on high-bandwidth routed connections

PIM-SM (sparse mode):

- *Pull* concept
- Does not forward multicasts unless requested to by a join message
- Periodic joins required to maintain the tree
- Used on lower bandwidth routed connections such as a WAN

A VRF can run PIM-DM or PIM-SM – not both

MOD 1- 11

The routing switches can construct the multicast routing tree in one of two modes:

PIM-Dense Mode (DM)

PIM-DM (RFC 3973) assumes everyone wants to receive the stream – a push concept. When a source begins streaming traffic, PIM-DM-enabled VLAN interfaces flood multicasts to all neighbors. This continues until neighbors send a prune message – “stop sending that stream to me”. Upstream neighbors set a timer for the pruned interface. When the timer expires, the interface reverts to the default behavior—forwarding multicasts until explicitly requested not to. You will soon learn how to prevent the timer from expiring.

This mode works best in environments with high bandwidth and high tolerance for congestion, such as a high-speed Ethernet network. Because the PIM-DM routers automatically create the multicast routing tree for a particular stream when that stream begins, PIM-DM is quite simple to set up.

PIM Sparse Mode (SM)

PIM-SM (RFC 4601) assumes nobody wants to receive the stream – more of a pull concept. It builds the multicast routing tree for a group in advance and then forwards multicasts on that tree when a source begins to stream to that group. PIM-SM-enabled routing interfaces must specifically indicate that they want to receive multicast traffic in a particular group by sending a

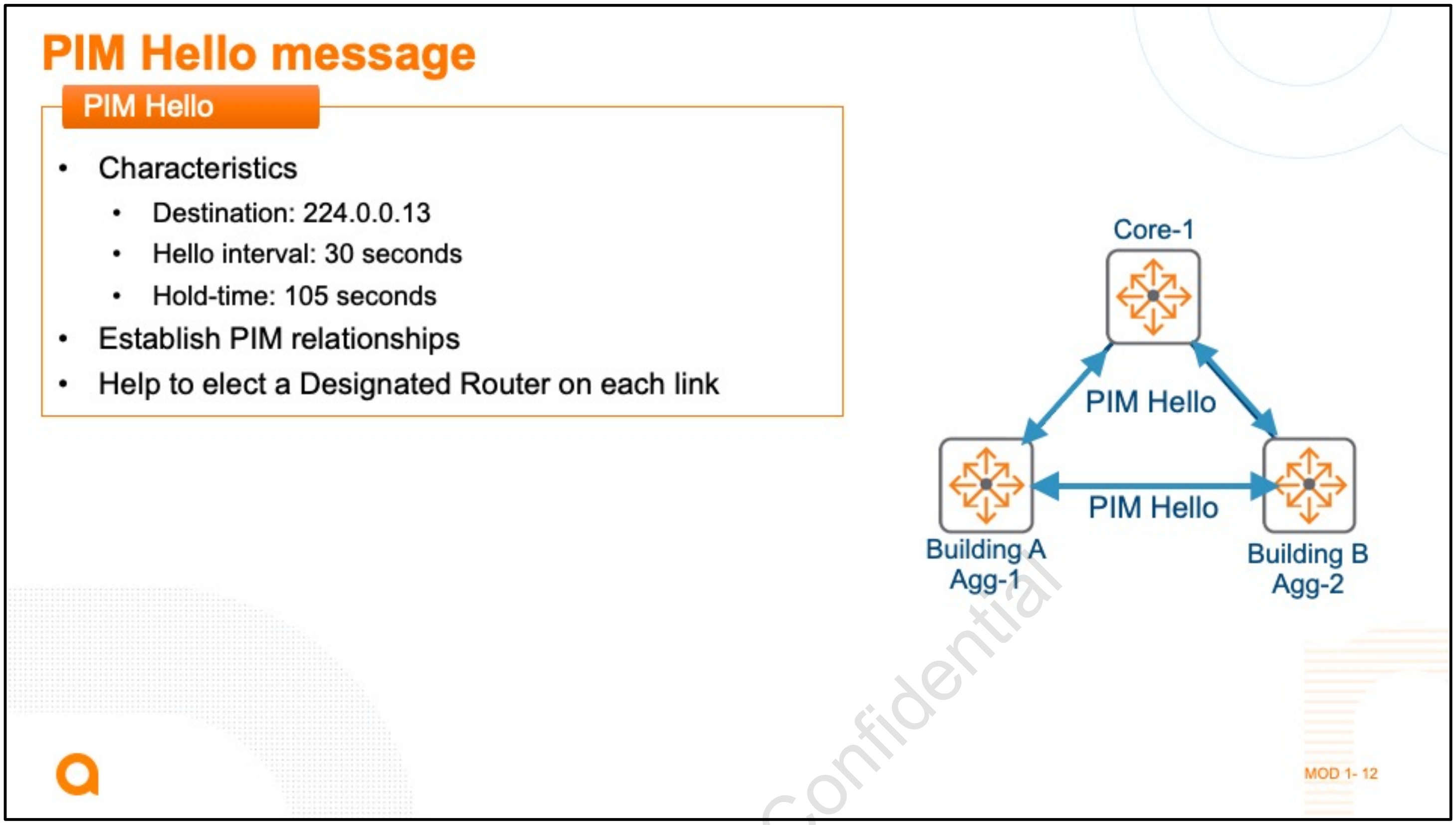
PIM join message. The interface must then periodically refresh its join. If its upstream neighbor does not receive joins, it reverts to its default behavior: not forwarding traffic.

Because PIM-SM establishes efficient forwarding paths, it is well-suited for a lower-bandwidth environment, such as one built on WAN links. PIM-SM also gives you more control over which multicast groups are routed.

PIM-SM has more configuration options and tends to be more complex to set up. PIM-DM can waste bandwidth during initial tree establishment, and may not be suitable for networks with many lower-speed WAN links.

Your AOS-CX switch supports both PIM-DM and PIM-SM; however it can only one run mode concurrently (within a VRF). All multicast routing switches in the network (in their respective VRF) must also agree and run the same type of PIM.

PIM-DM interoperates with IGMP/MLD and the switch's routing protocols. It does not matter which unicast routing protocol is in use – PIM can use it for RFP. That is why it is called the Protocol Independent Multicast routing protocol. PIM-DM can be used with RIP, OSPF, BGP, or static routes configured.



PIM Hello message

PIM Hello

- Characteristics
  - Destination: 224.0.0.13
  - Hello interval: 30 seconds
  - Hold-time: 105 seconds
- Establish PIM relationships
- Help to elect a Designated Router on each link

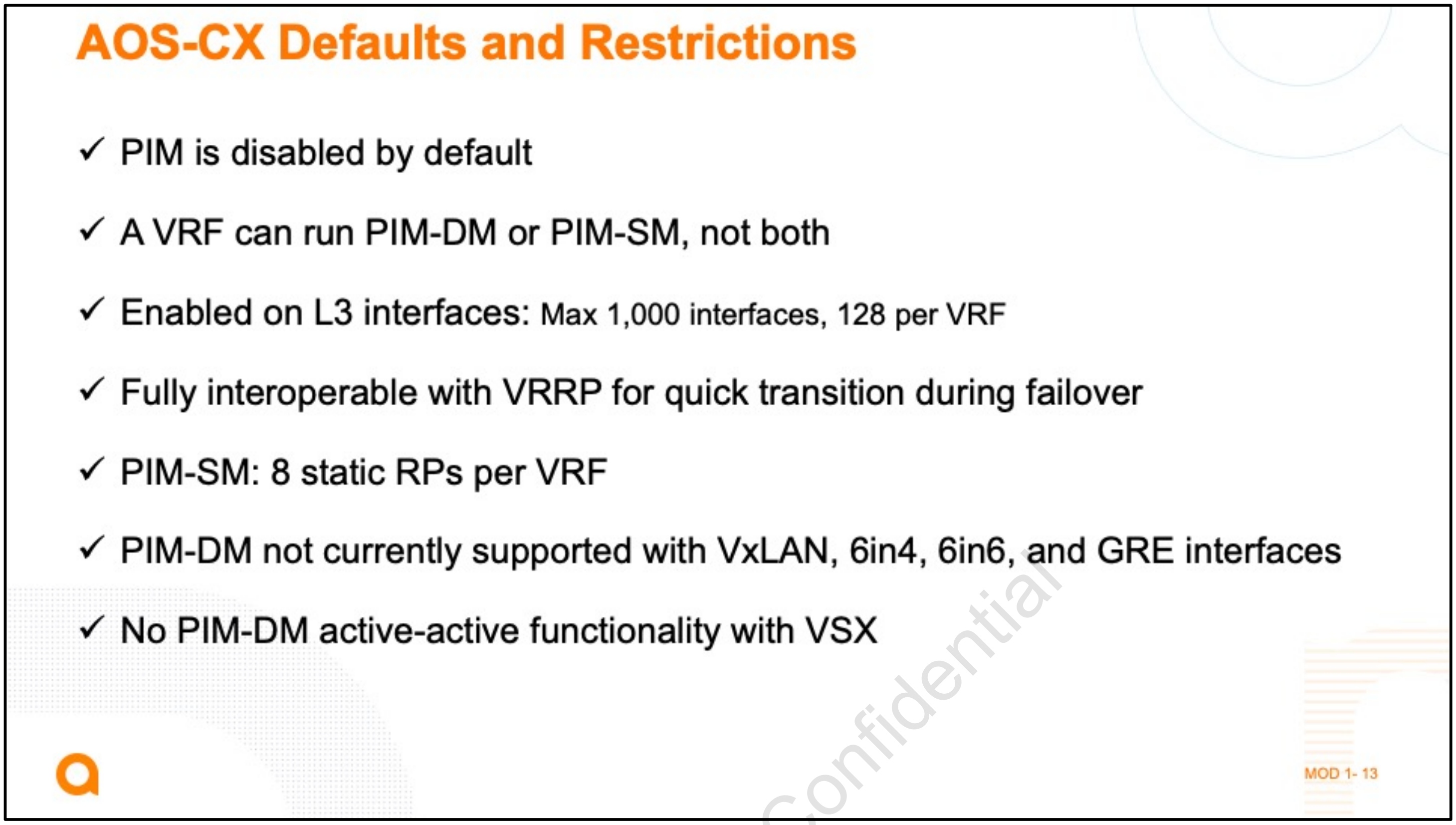
MOD 1- 12

Routers and L3 switches running PIM send periodic hello messages on PIM-enabled interfaces to discover neighbors. These messages are sourced with the IP address of the outgoing interface (In ArubaOS-CX, you can specify the interface that should be used to source those packets) and are sent to all PIM routers with the multicast destination address of 224.0.0.13. They contain information such as the hello period time, hello hold-time, LAN prune delay time and capabilities. The hold-time signifies how long the neighbor will keep the neighbor relationship without receiving an update. The LAN prune delay is used on multi-access networks for propagation delay.

The packet capture in the figure shows a hello message. Pay special attention to the highlighted areas - the destination multicast addresses at Layer 3, the PIM version (v2) and the PIM message type (Hello).

The PIM process begins when two or more routers establish a PIM neighbor adjacency on a given segment. To begin the adjacency process, any interfaces participating in multicast routing must be configured for PIM communications. Routers discover PIM neighbors by sending PIM hello messages to the link-local multicast address 224.0.0.13 out each PIM-enabled interface.

Hello messages are used to elect a Designated Router - a single PIM router that acts on behalf of directly connected hosts in a shared media, like Ethernet. Hello messages are also the way that option negotiation takes place in PIM, so that additional functionality can be enabled, or parameters tuned.



AOS-CX Defaults and Restrictions

- ✓ PIM is disabled by default
- ✓ A VRF can run PIM-DM or PIM-SM, not both
- ✓ Enabled on L3 interfaces: Max 1,000 interfaces, 128 per VRF
- ✓ Fully interoperable with VRRP for quick transition during failover
- ✓ PIM-SM: 8 static RPs per VRF
- ✓ PIM-DM not currently supported with VxLAN, 6in4, 6in6, and GRE interfaces
- ✓ No PIM-DM active-active functionality with VSX

MOD 1- 13

PIM (DM or SM) is disabled by default. Either PIM-SM or PIM-DM can be configured within a VRF at a time. All the interfaces within the VRF must run with same mode.

PIM uses unicast routing information from any of the routing protocols that are running on the system, such as OSPFv2, OSPFv3, BGP. Static routes are also supported with next-hop IP addresses.

PIM can be enabled on an ROP (routed port) and SVI (interface vlan) interfaces. PIM can be enabled across all VRFs on a maximum of 1,000 interfaces with an upper limit of 128 per VRF.

Note: Although up to 128 PIM DM enabled interfaces can be configured, when configuring trunk interfaces with multiple Dense enabled SVIs, the trunk interfaces must have sufficient bandwidth or have only the required number of trunks it can support. This ensures that the link utilization is not exceeded due to the initial flooding nature of the protocol.

PIM-DM is compatible with IGMP version 2 and version 3, MLD version 1 and version 2, and is fully interoperable with IGMP/MLD for determining multicast flows. PIM-DM is fully interoperable with VRRP to quickly transition multicast routes in a failover. PIM-DM can run on multiple VRF instances in parallel. It is supported on all VRFs supported in the system.

PIM-DM currently does not support the following:

VxLAN, 6in4, 6in6, and GRE interfaces

PIM-DM can be enabled on VSX deployments, however active-active functionality is not available.

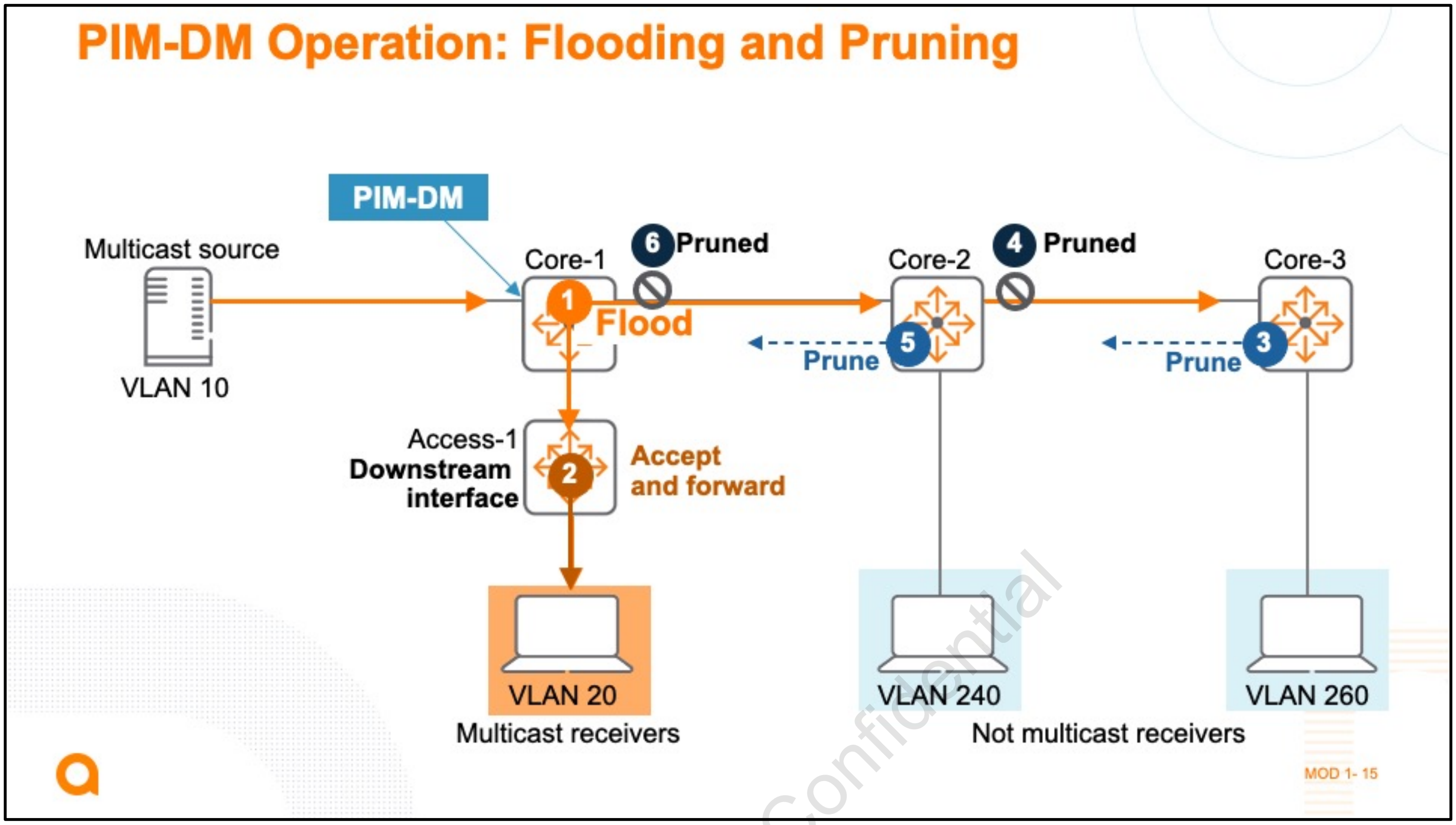
Note: Since active-active is not supported a slightly higher traffic recovery time can be expected compared to sparse mode, in the event of failovers. Hence, it is recommended to use PIM-SM on VSX. Ensure that the ISL is not oversubscribed while using PIM-DM with VSX. In case of numerous PIM enabled SVIs, ensure that ISL is configured to handle the flooding on all VLANs since ISL is member of all SVIs.

If BFD is enabled globally, it will be enabled by default on all PIM interfaces as well. The only exception is when it is disabled specifically on an interface using the ip pim-dense bfd disable interface command. If BFD is disabled globally, it will be disabled by default on all interfaces. The only exception is when it is enabled specifically on an interface using the ip pim-dense bfd interface command.

PIM-DM

MOD 1- 14

PIM Dense Mode is a great option for many deployments. Let’s take a look.



PIM-DM Operation: Flooding and Pruning

A multicast source begins to transmit, and its default router receives the multicasts. The VLAN interface on which the multicasts arrive should be enabled for PIM-DM. The router then floods the multicasts on every PIM-DM interface, including:

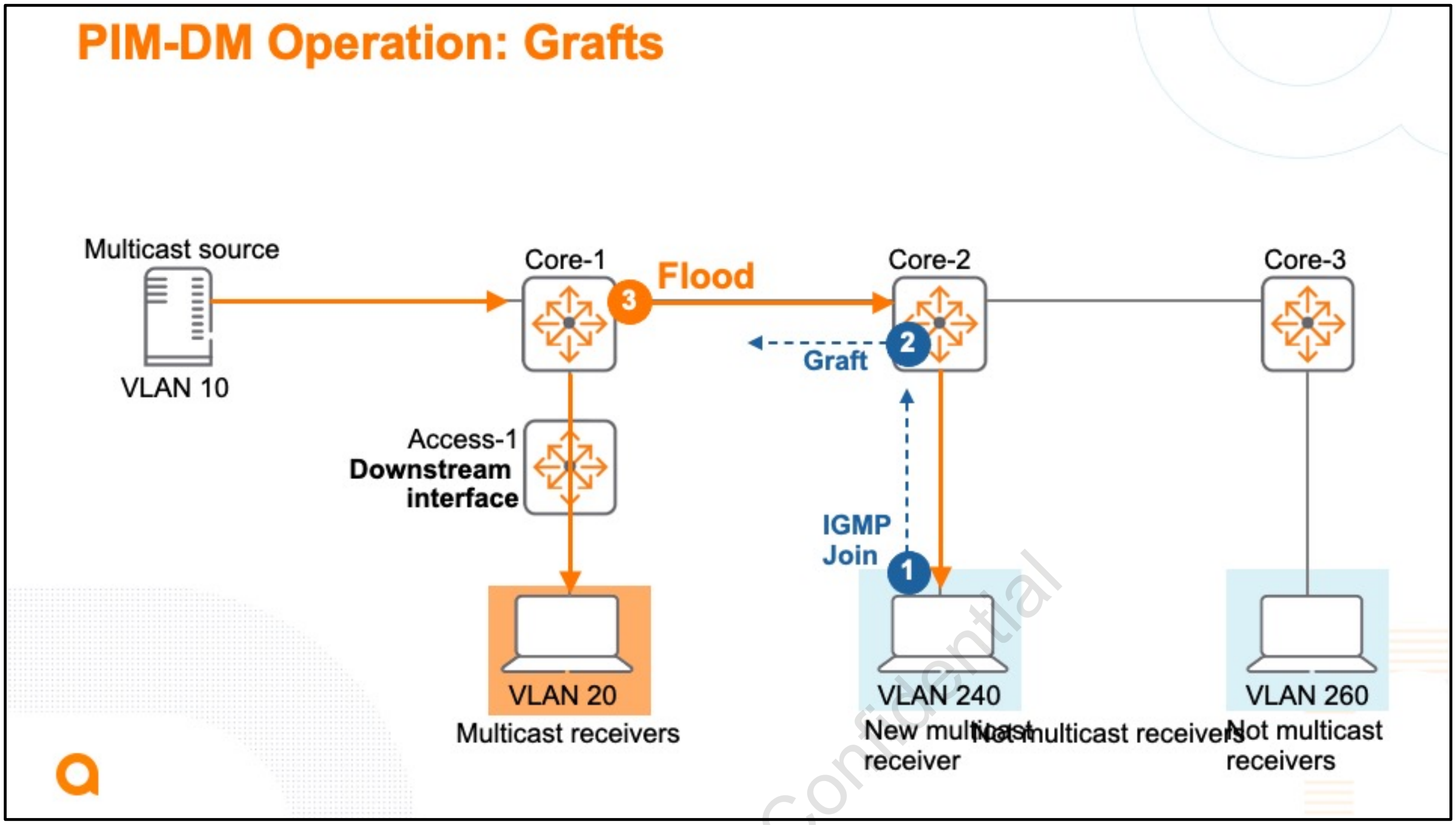
Interfaces that connect to PIM-DM neighbors - Core-2 and Access-1 in the figure.

Interfaces that connect to receivers that have requested the multicast using IGMP. In this example, Core-1 has no directly attached multicast receivers

Downstream multicast routers follow the same steps to flood the multicast. In figure, Access-1 and Core-2 are only neighbors with Core-1. Access-1 has learned that VLAN 20 is a downstream interface for this group from IGMP, and so continues to accept and forward traffic from Core-1.

Neither Core-2 or Core-3 have multicast receivers, so they send prune messages to stop the stream from flowing to them. Core-2 still maintains the multicast route [the (S, G) entry), which it can use to receive multicasts later, if necessary.

To summarize, for a multicast group "X" on a given interface, when the last host belonging to group "X" leaves the group, PIM places that interface in a prune state. Multicast traffic from group "X" is now blocked to that interface. The prune state remains until a host on the same interface issues a join for group "X", in which case the router cancels the prune state and changes the flow to the forwarding state.



PIM-DM Operation: Grafts

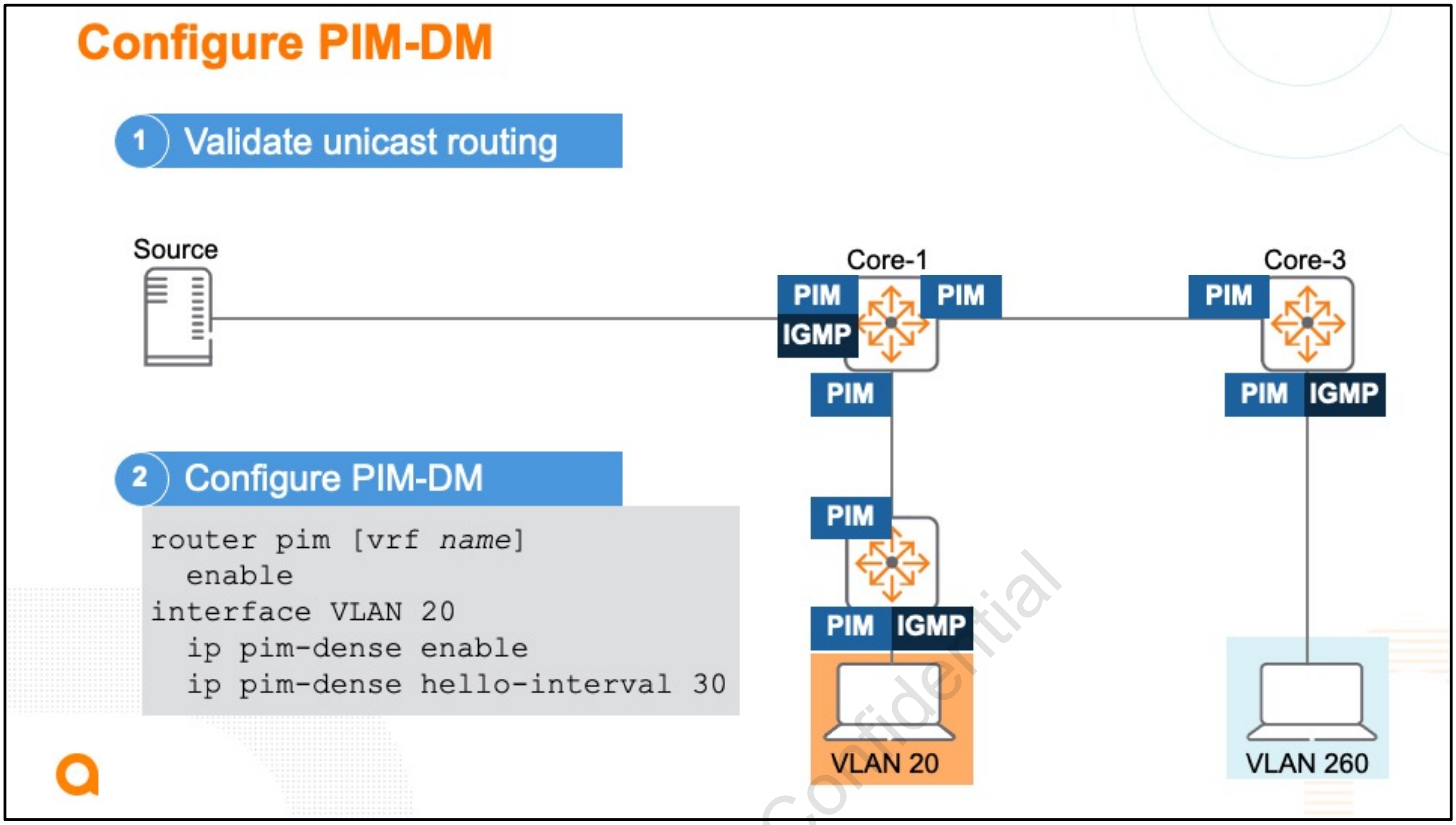
After pruning, Core-1 only forwards multicast traffic to Access-1, which forwards the traffic to receivers in VLAN 20.

As group memberships change, the Shortest Path Tree (SPT) must change to accommodate. Recall that PIM-DM prefers to forward unnecessary multicast traffic over failing to forward necessary traffic. Therefore, prune messages expire, and the routing switch that had received the prune message begins forwarding multicast traffic again.

In addition, group members leave groups. For example, if all multicast receivers in VLAN 20 left the group, then Access-1 would send a prune message to Core-1, stopping the flow of traffic.

The figure above illustrates a new host that wishes to receive the multicast stream, probably because some user sat at this host and opened an application that required the stream. The host sends an IGMP report or join message for this group, and Core-2 adds VLAN 240 as a downstream interface for its (S, G) multicast route for this multicast stream. Eventually, the prune message would expire, and Core-2 would receive traffic again. However, PIM-DM provides a faster method of pulling the multicast traffic down. Core-2 sends a graft message on the upstream interface, requesting to be added to the tree.

Core-1 adds the interface that connects to Core-2 back to the (S, G) multicast route's list of downstream interfaces. Core-2 begins to receive the traffic again, and it forwards it to the receiver.



Configure PIM-DM

1 Validate unicast routing

2 Configure PIM-DM

```
router pim [vrf name]
  enable
interface VLAN 20
  ip pim-dense enable
  ip pim-dense hello-interval 30
```

PIM-DM must be enabled on all VLAN interfaces between the multicast sources and the multicast receivers:

Interfaces on which multicast sources reside

Interfaces on which multicast receivers reside (IGMP-enabled interfaces)

Interfaces between all routers and routing switches that connect sources and receivers

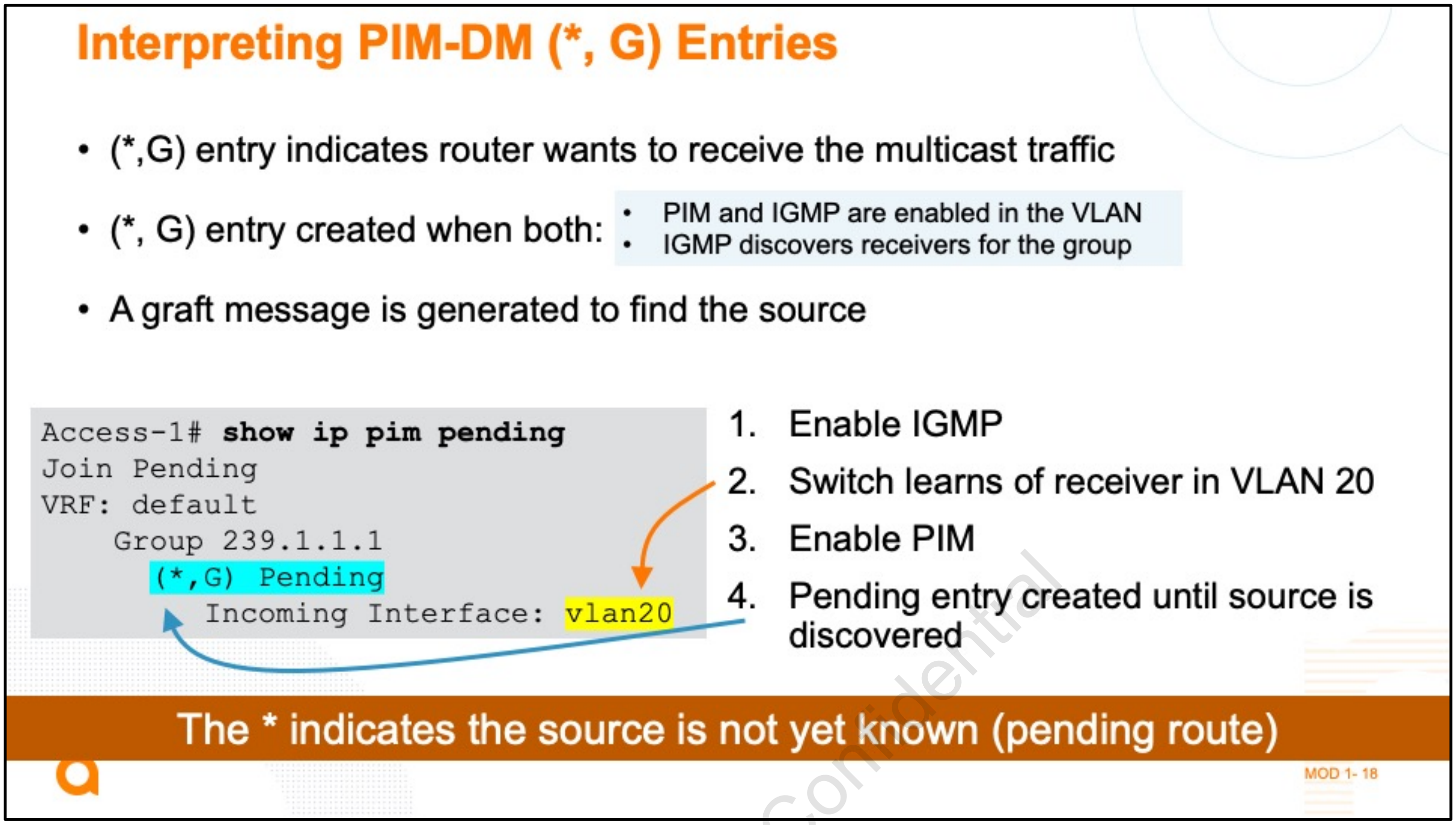
The figure shows where to enable PIM-DM and IGMP. Note that you should also enable IGMP on the server VLAN; although IGMP is not strictly required in the source VLAN, it enables Core-1 to filter multicasts within that VLAN.

First validate that unicast routing is in place and fully functional. Next, enable PIM-DM multicast routing globally. Then enable PIM-DM on appropriate VLAN interfaces, as shown in the figure.

A PIM router uses hello packets to automatically discover neighbors. They are sent to 224.0.0.13, every 30 seconds by default. You can change the default within the range of 5-300 seconds. The router uses this setting to compute the hello hold-time, which is included in hello packets sent to neighbor routers. The hello hold-time tells neighbor routers how long to wait for the next hello packet from the routing switch. If another packet does not arrive within that time, the router removes the neighbor adjacency on that interface from the PIM adjacency table, which removes any flows running on that interface. Shortening the hello interval reduces the hello hold-time. If they do not receive a new hello packet when expected, it changes how quickly other routers stop sending traffic to the routing switch.

BFD can be used to speed up the convergence of PIM. If BFD is globally enabled, BFD is automatically enabled for PIM on interfaces where PIM is also enabled. BFD was previously covered.

Note: MLD and PM-DM are not currently supported for IPv6 in AOS-CX; however MLD and PIM-SM are.



Interpreting PIM-DM (*, G) Entries

- (*,G) entry indicates router wants to receive the multicast traffic
- (*, G) entry created when both:
  - PIM and IGMP are enabled in the VLAN
  - IGMP discovers receivers for the group
- A graft message is generated to find the source

```
Access-1# show ip pim pending
Join Pending
VRF: default
    Group 239.1.1.1
      (*,G) Pending
         Incoming Interface: vlan20
```

1. Enable IGMP
2. Switch learns of receiver in VLAN 20
3. Enable PIM
4. Pending entry created until source is discovered

The * indicates the source is not yet known (pending route)

MOD 1- 18

When you are monitoring and troubleshooting a multicast routing solution, you do not have an illustration tracing multicasts for you. It can be helpful for you to understand more about the multicast route entries, how to interpret them, and how the routing switch uses them to forward traffic.

Begin with (*, G) entries.

When you enable PIM-DM on a VLAN interface, the switch creates a (*, G) entry for each group membership discovered by IGMP on that VLAN. This entry specifies the downstream interfaces from which a tree can be formed when multicasts begin to arrive from a particular source.

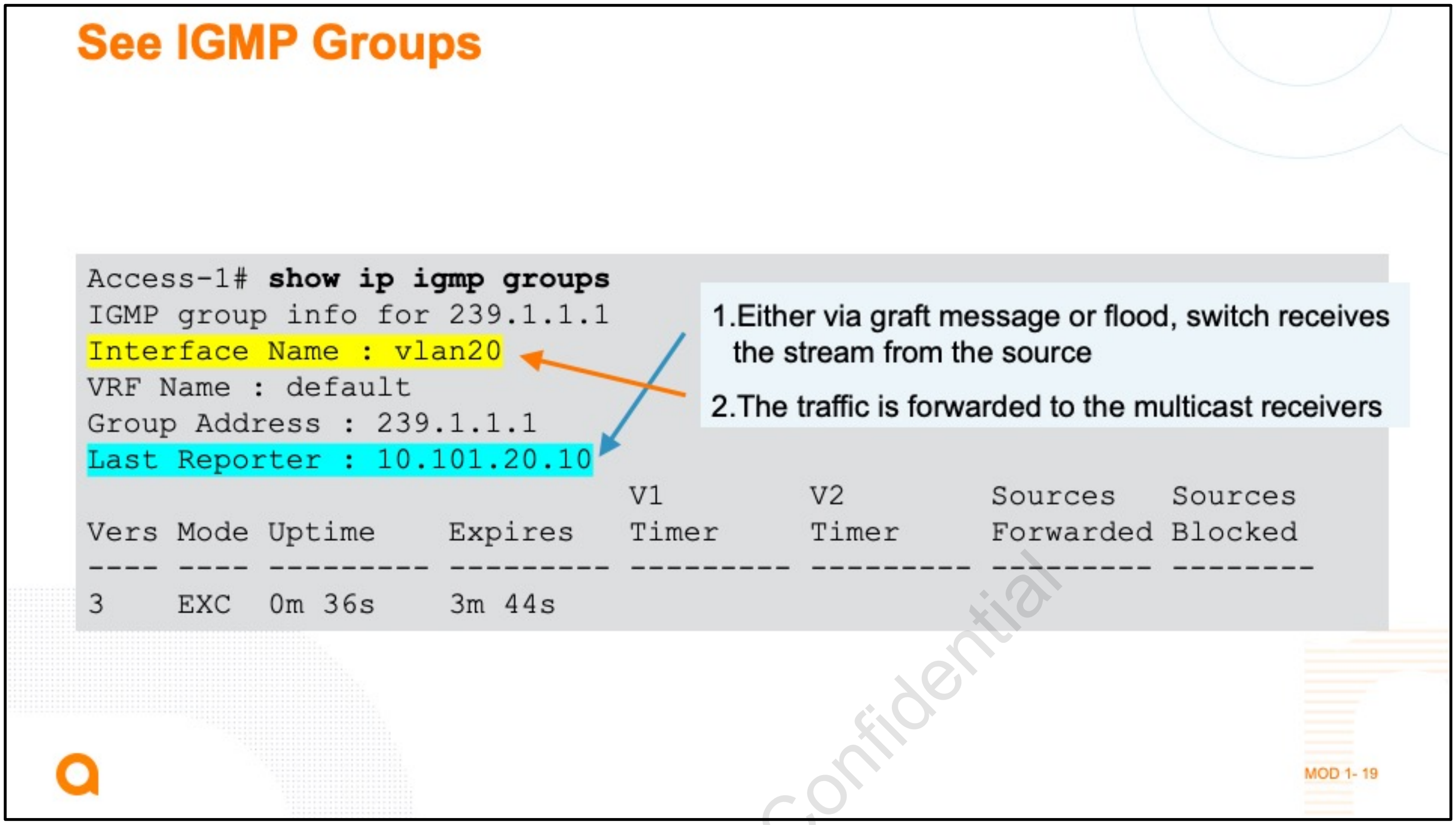
When you use PIM-DM, the (*, G) entry essentially acts as a placeholder, listing the downstream interfaces that require multicasts in this group, as discovered by IGMP. The framework of the tree that delivers these multicasts is not filled in. For example, no upstream interface is specified (the routing switch does not know where multicasts will arrive).

The (*, G) entry exists only as a pending route; it is not an active route forwarding multicasts. The (*, G) entry remains in the PIM routing table for as long as an IGMP membership for the group exists on at least one interface. This signals to the routing switch that it should stay part of the multicast routing tree for any multicast stream to this group, regardless of whether it needs to forward the traffic to any downstream PIM-DM neighbors.

Typically you'll see (*,G) entries when there has been a purge. Now a client wants to receive a stream, but the switch isn't receiving the stream because of a previous purge. The switch generates a graft message and sends it to its upstream neighbors to find the source so that traffic can be flooded back to the switch.

You also see (*,G) entries if there is a network issue where the source is no longer reachable and thus no multicast stream is being received, yet the switch still has clients locally or downstream that desire to receive the stream.

To view (*,G) routes, enter the show ip pim pending command. To view information about IGMP groups, enter show ip igmp groups or show ip igmp group command, qualifying it with a specific multicast address.



See IGMP Groups

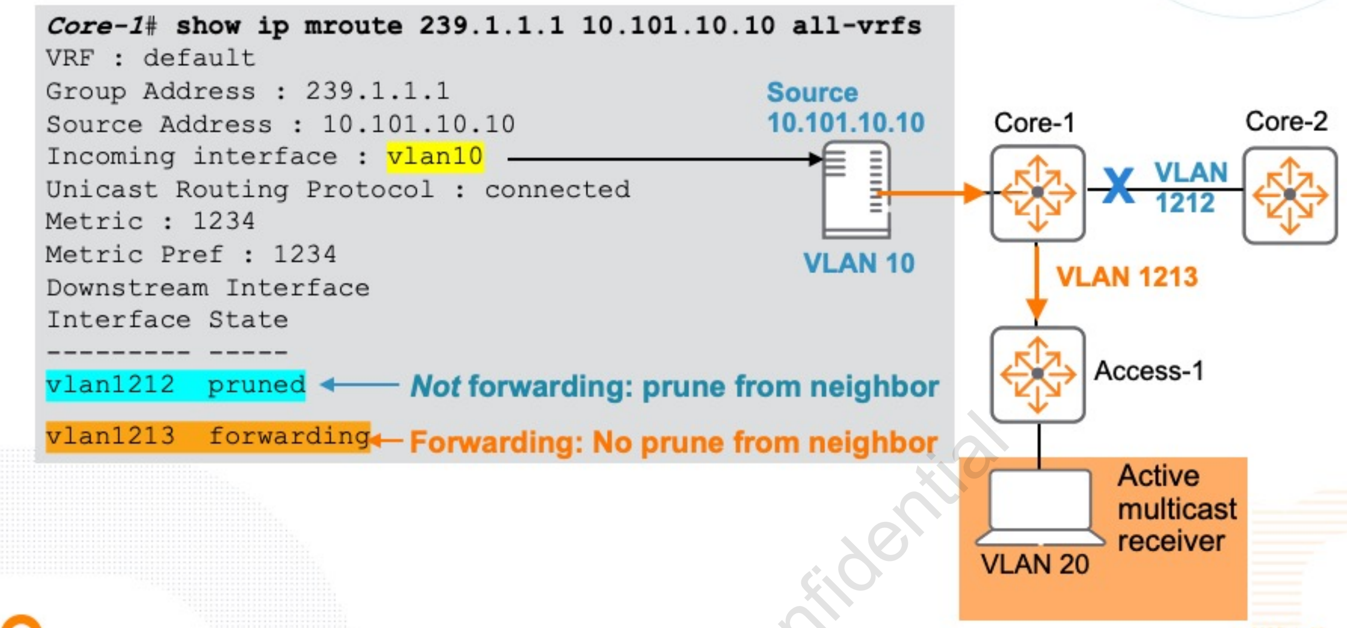
```
Access-1# show ip igmp groups
IGMP group info for 239.1.1.1
Interface Name : vlan20
VRF Name : default
Group Address : 239.1.1.1
Last Reporter : 10.101.20.10
                              V1        V2        Sources   Sources
Vers Mode Uptime    Expires   Timer     Timer     Forwarded Blocked
---- ---- --------- --------- --------- --------- --------- --------
3    EXC  0m 36s    3m 44s
```

1. Either via graft message or flood, switch receives the stream from the source
2. The traffic is forwarded to the multicast receivers

MOD 1- 19

To view (*,G) routes, enter the show ip pim pending command. To view information about IGMP groups, enter show ip igmp groups or show ip igmp group command, qualifying it with a specific multicast address

See IGMP Routes Pruned



```
Core-1# show ip mroute 239.1.1.1 10.101.10.10 all-vrfs
VRF : default
Group Address : 239.1.1.1
Source Address : 10.101.10.10
Incoming interface : vlan10
Unicast Routing Protocol : connected
Metric : 1234
Metric Pref : 1234
Downstream Interface
Interface State
--------- -----
vlan1212  pruned
vlan1213  forwarding
```

Not forwarding: prune from neighbor

Forwarding: No prune from neighbor

MOD 1- 20

When the switch receives a multicast on the proper upstream interface for the source, it creates an (S, G) entry. It uses this entry to route the multicast traffic on the correct interface or interfaces.

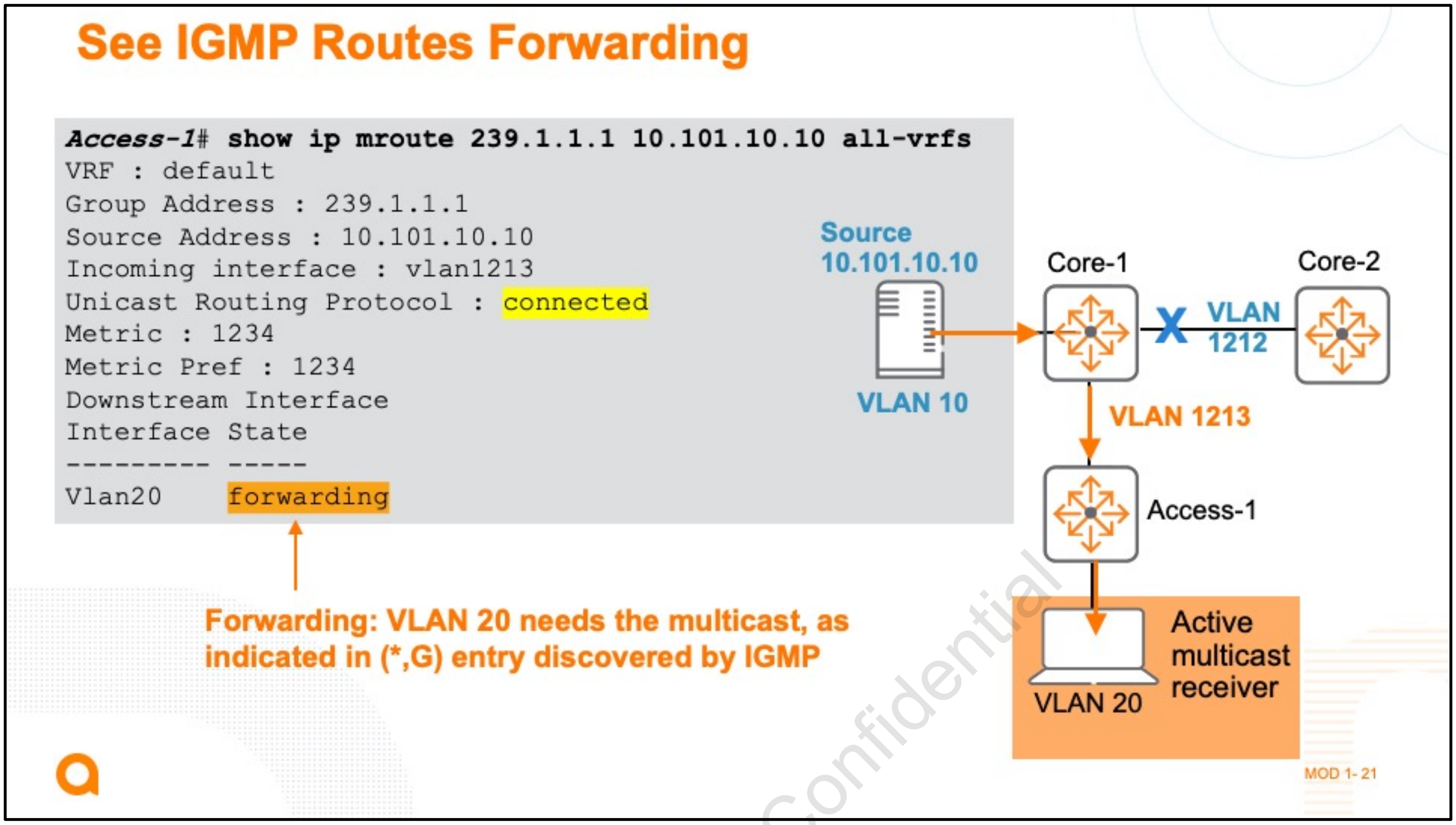
The (S, G) entry includes:

The source for the incoming multicasts

The upstream interface on which the multicast arrived
You will learn in a moment how the switch checks the validity of this interface. It will not add the route if the upstream interface is invalid.

The downstream interfaces
Initially, the downstream interface includes all PIM-DM interfaces set to a forwarding state. The switch maintains the forwarding state unless a neighbor sends a prune on that interface. The switch also always maintains the forwarding state for interfaces discovered by IGMP as long as they are in the (*,G) entry associated with this group.

As you see, if the interface receives a prune, the switch maintains the interface in the list, but places it in the pruned state and does not forward multicasts on it.

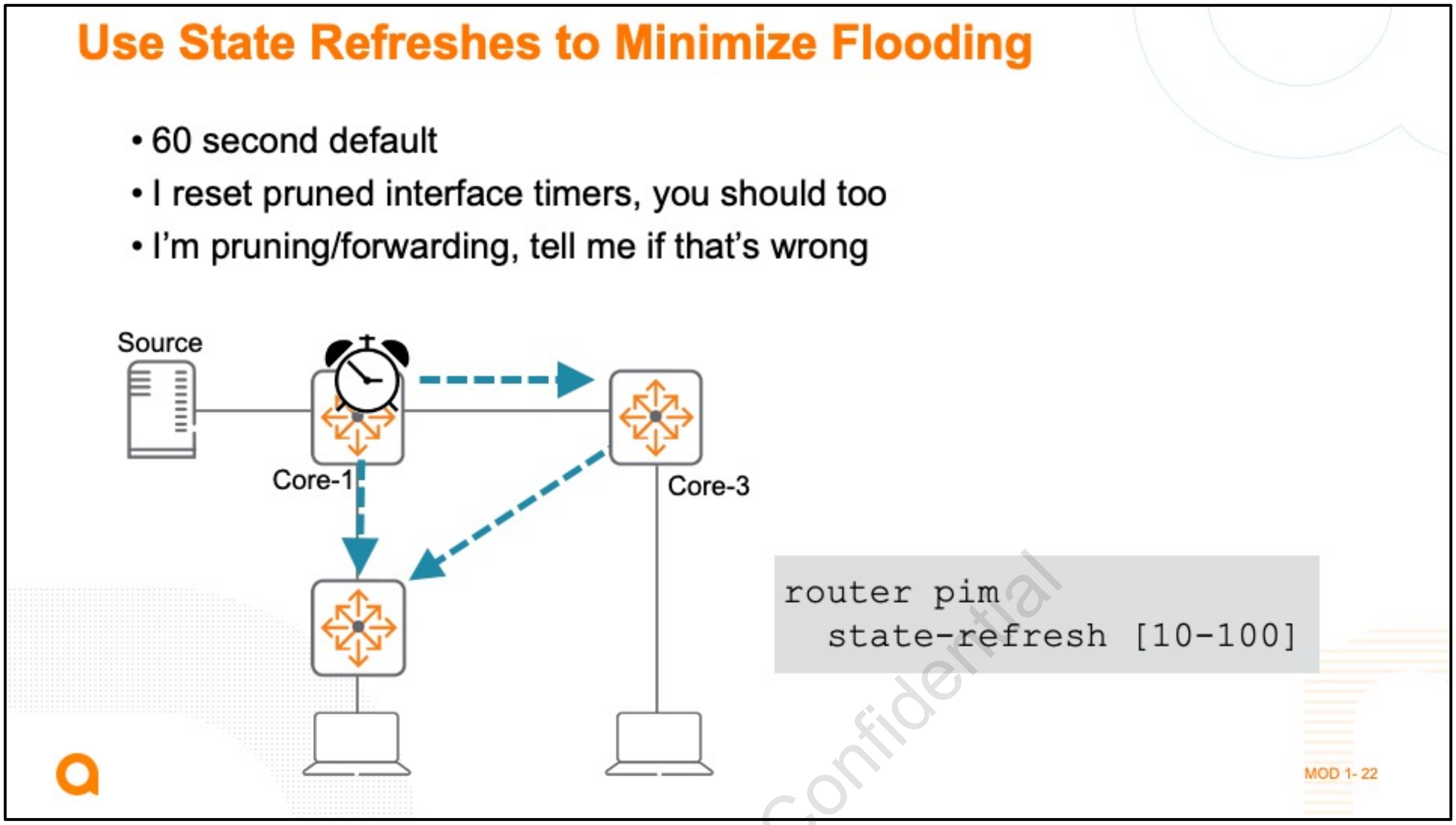


See IGMP Routes Forwarding

```
Access-1# show ip mroute 239.1.1.1 10.101.10.10 all-vrfs
VRF : default
Group Address : 239.1.1.1
Source Address : 10.101.10.10
Incoming interface : vlan1213
Unicast Routing Protocol : connected
Metric : 1234
Metric Pref : 1234
Downstream Interface
Interface State
--------- -----
Vlan20    forwarding
```

Forwarding: VLAN 20 needs the multicast, as indicated in (*,G) entry discovered by IGMP

If the (S, G) entry has downstream interfaces or if it the source is connected directly to the routing switch, the routing switch adds the (S, G) entry to its active multicast routing table. The routing switch then uses the (S, G) entry to forward multicasts.



Use State Refreshes to Minimize Flooding

- 60 second default
- I reset pruned interface timers, you should too
- I'm pruning/forwarding, tell me if that's wrong

```
router pim
  state-refresh [10-100]
```

MOD 1- 22

You learned that prunes periodically time out, causing PIM-DM-enabled interfaces to resume the multicast flood to previously pruned neighbors. If the neighbor still does not require the traffic, it must send another prune message. These bursts of unnecessary traffic can contribute to congestion.

AOS-CX switches can automatically prevent the prune state time out - the source-connected switch floods state refresh messages on all PIM-DM interfaces. This indicates state information about the transmitting interface, whether pruned or forwarding. It propagates throughout the system, triggering all PIM-DM routers to send joins and prunes only if they need to change the current state of the upstream interface. Routers can then also refresh their prune states with confidence without needing to flood multicast traffic.

In more detail, receiving a state refresh message triggers the PIM-DM interface to take these actions:

Drop the message unless it has arrived from the RPF upstream neighbor.

For messages from the proper upstream neighbor:

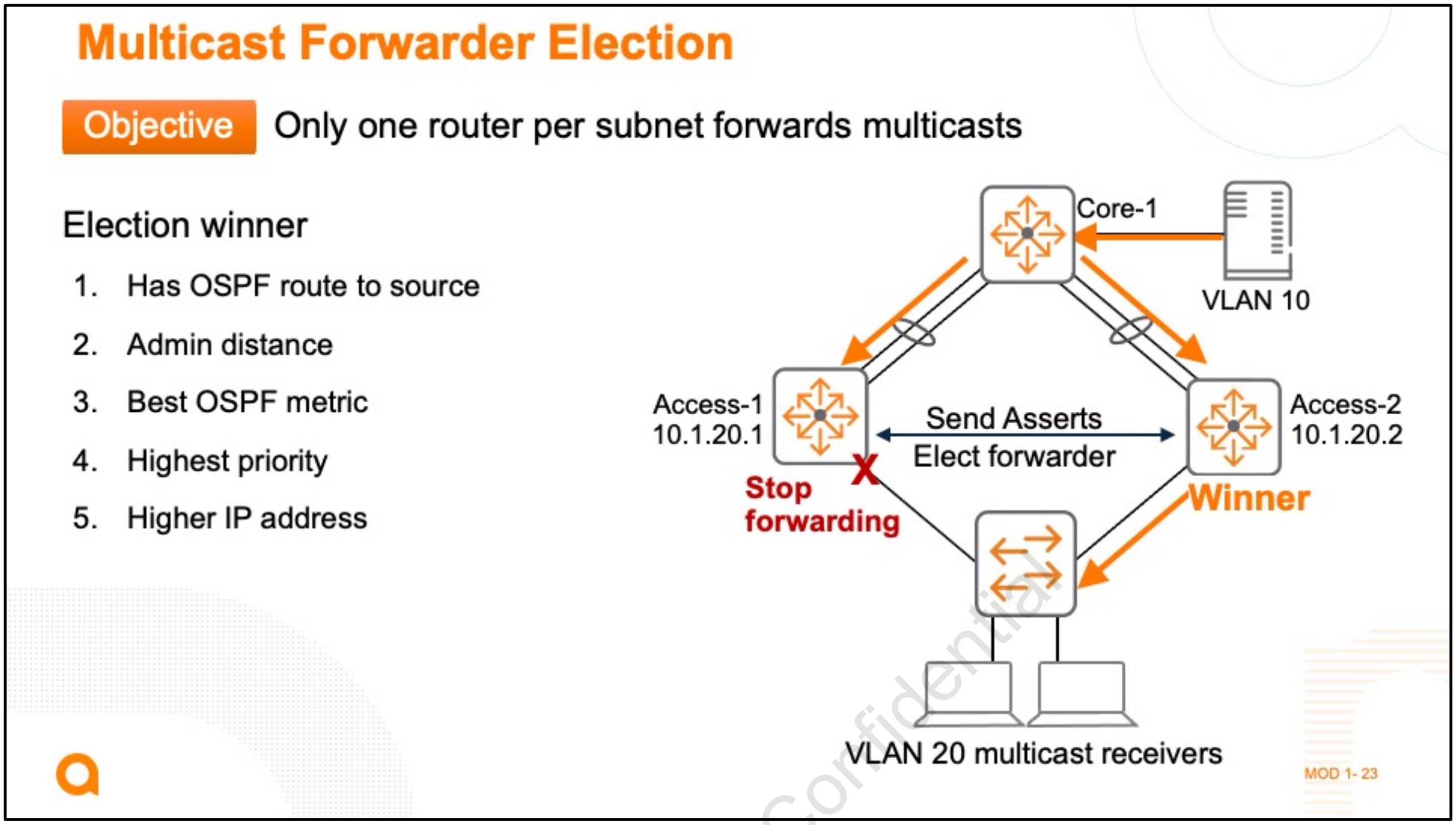
- Check the indicated state of the upstream interface and determine whether it matches this routing switch's requirements.
- If the indicated state is pruned and this routing switch has no downstream interfaces for the (S, G) entry, the indicated state is correct. Similarly, if the indicated state is forwarding and this routing switch does have downstream interfaces on which it must forward traffic, the

state is correct.

- If the indicated state is pruned, and this switch requires the multicasts, the switch sends an (S, G) join upstream to alter the state. Similarly, if the indicated state is forwarding, but this switch does not have an active entry forwarding multicast traffic downstream, the switch sends an (S, G) prune upstream.
- Reset all pruned interface timers.
  The switch can safely reset all timers because its downstream neighbors will correct any mistakes on receipt of the state refresh message, just as this routing switch did.
- Forward the state refresh message (with one less value in the TTL) on all interfaces with PIM-DM neighbors except the interface on which it received the message.
  The TTL prevents the state refresh message from endlessly looping through the PIM-DM domain.

AOS-CX switches automatically support state refresh messages, but only the routing switch connected directly to the unicast source initiates state refresh messages. All other PIM routers in the network only propagate these state-refresh messages.

You can configure the interval in seconds (10 to 100) between successive state refresh messages. The default setting is 60 seconds. You should keep the interval below the PIM-DM prune expiry time to prevent periodic unnecessary floods of traffic.



Multicast Forwarder Election

Objective Only one router per subnet forwards multicasts

Election winner

1. Has OSPF route to source
2. Admin distance
3. Best OSPF metric
4. Highest priority
5. Higher IP address

MOD 1- 23

In this scenario, routers Access-1 and Access-2 have interfaces in VLAN 20, connected to a downstream switch with multicast receivers. It has chosen this solution rather than a VSF fabric, which is typically recommended. The switches run VRRP and PIM-DM. Just as only one of the redundant switches can act as the active VRRP master, only one of the switches can forward multicasts. Otherwise, endpoints could receive duplicate copies of the multicasts, causing errors.

When Access-1 and Access-2 receive multicasts from Core-2, they flood the multicasts to each other. As you learned earlier, Reverse Path Forwarding (RPF) ensures that only the best path continues to forward frames. Let’s look at this in more detail.

The PIM routers send an assert message to elect a forwarder, then examine each other’s asserts to elect a multicast forwarder for the interface. The assert includes the (S, G) entry under dispute and information about the PIM routing switch’s unicast path to the source. The routing switch with the more favorable path becomes the multicast forwarder, as determined first by administrative distance (preference) and then, if the administrative distance ties, by metric or cost. If the metric is the same, the router with the highest priority value is elected. If that matches, the interface with the higher IP address is elected.

Suppose that both routers have the same admin distance, cost, and priority for the route. However, Access-2 has the higher IP address. Therefore, Access-2 is elected multicast forwarder for VLAN 20.

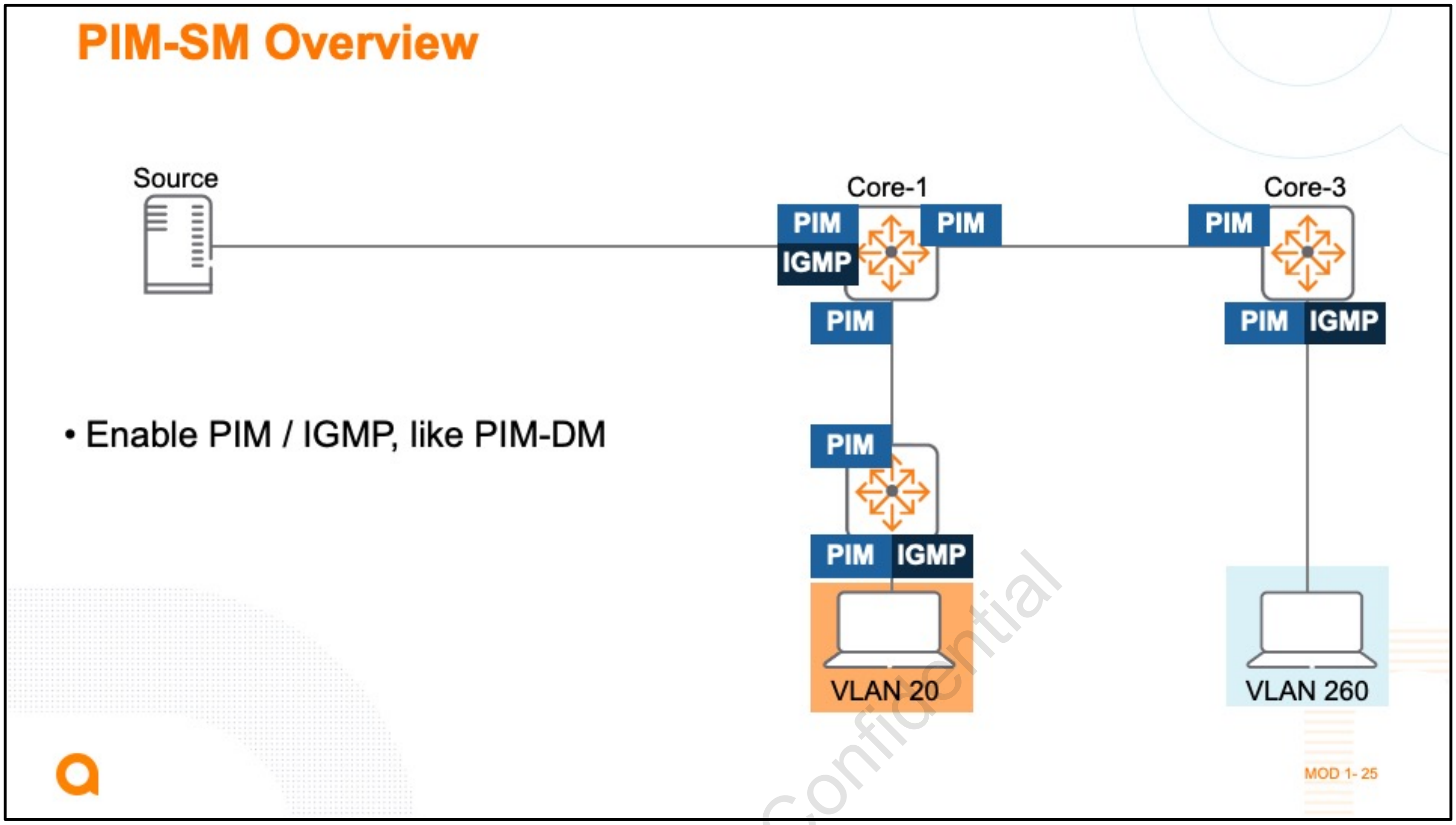
Access-1 removes this VLAN from its list of downstream interfaces in its (S, G) entry. It also sends a prune to Access-2. Access-2 continues to forward multicasts in VLAN 20 because multicast receivers connect in this VLAN. Of course, if all receivers send an IGMP leave message, there would be no receivers in VLAN 20. In that case Access-2 would send a prune message upstream to stop the unneeded multicast flow.

In a design like this you should pay attention to the IP addresses that you assign to routers. To load balance the Master and multicast forwarder role on different VLANs, assign one switch a higher IP address and VRRP priority on some VLANs and the other switch a higher IP address and VRRP priority on other VLANs.

PIM-SM

MOD 1- 24

You know Dense Mode, here’s the Sparse Mode Sibling.



PIM-SM Overview

- Enable PIM / IGMP, like PIM-DM

PIM-SM involves a little more setup than PIM-DM because PIM-SM routing switches cannot assume that they are part of the initial multicast routing tree for all traffic. Instead, PIM-SM routing switches must explicitly join that tree, and you must configure them to do so.

Recall that PIM-SM assumes that most hosts do not want to receive multicast traffic. It uses a non-flooding multicast model to direct traffic from the source to the interface when there are multicast receivers in the group. As a result, this model sends traffic only to the routers that specifically request it.

PIM-SM can be configured on physical ports, VLAN interfaces, LAG interfaces, and loopback interfaces.

Configuring PIM-SM is like configuring PIM-DM in some ways. For example, you must enable PIM-SM on every interface between multicast sources and receivers, including:

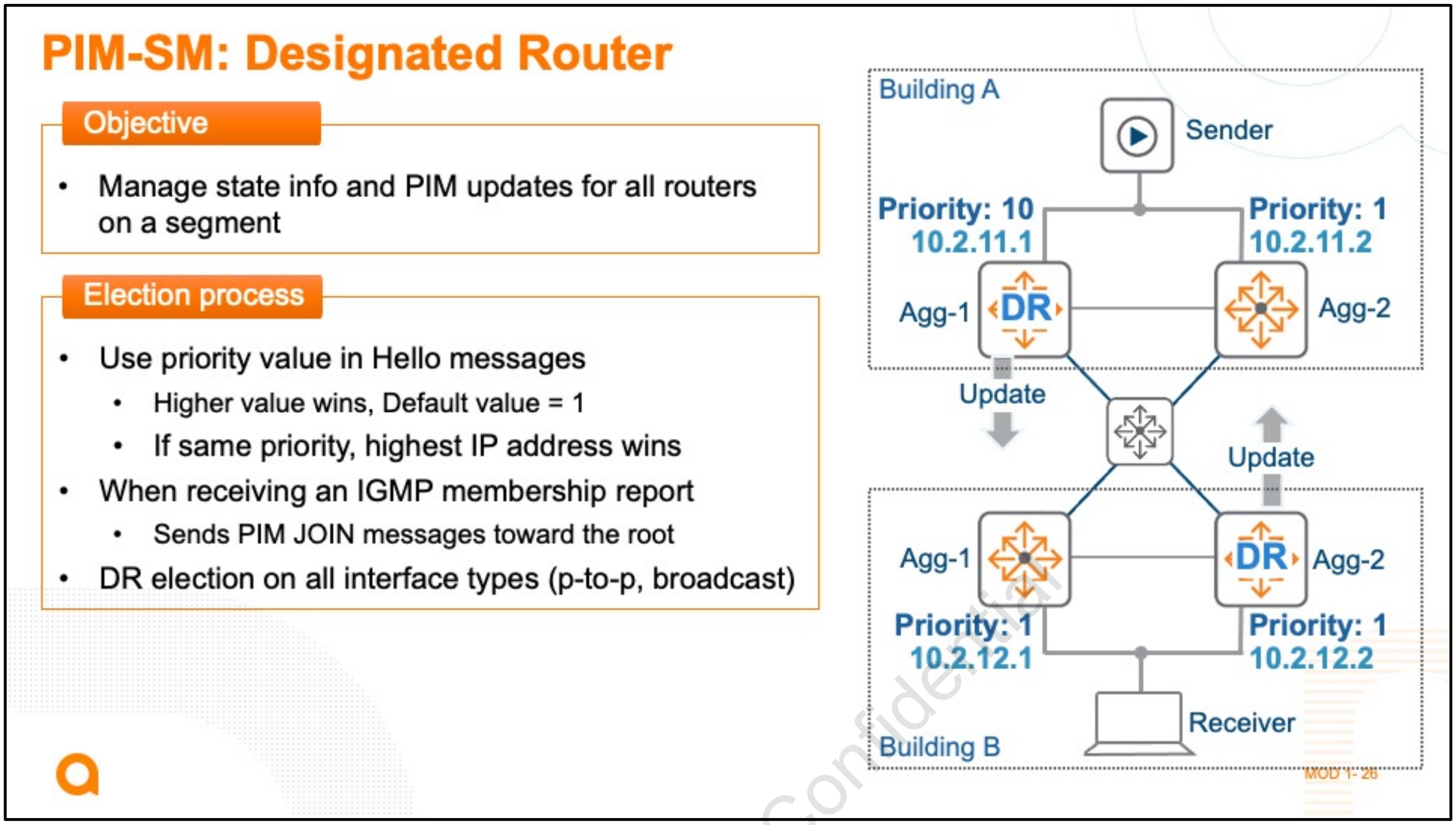
Layer 3 interfaces on which multicast sources reside

Layer 3 interfaces on which multicast receivers reside (IGMP-enabled interfaces)

Layer 3 interfaces between all routers and routing switches that connect sources and receivers

As with PIM-DM, PIM SM-enabled interfaces exchange hellos on 224.0.0.13. This helps explain why you cannot implement PIM-SM and PIM-DM in the same domain. PIM-enabled routing switches operating in one mode would expect certain behavior from neighbors operating in another, leading to failed establishment of the multicast forwarding tree.

In addition to establishing neighbor relationships, PIM interfaces in broadcast or multi-access networks elect a Designated Router (DR) - responsible for forwarding multicasts. Let’s look at that.



PIM-SM: Designated Router

Objective

- Manage state info and PIM updates for all routers on a segment

Election process

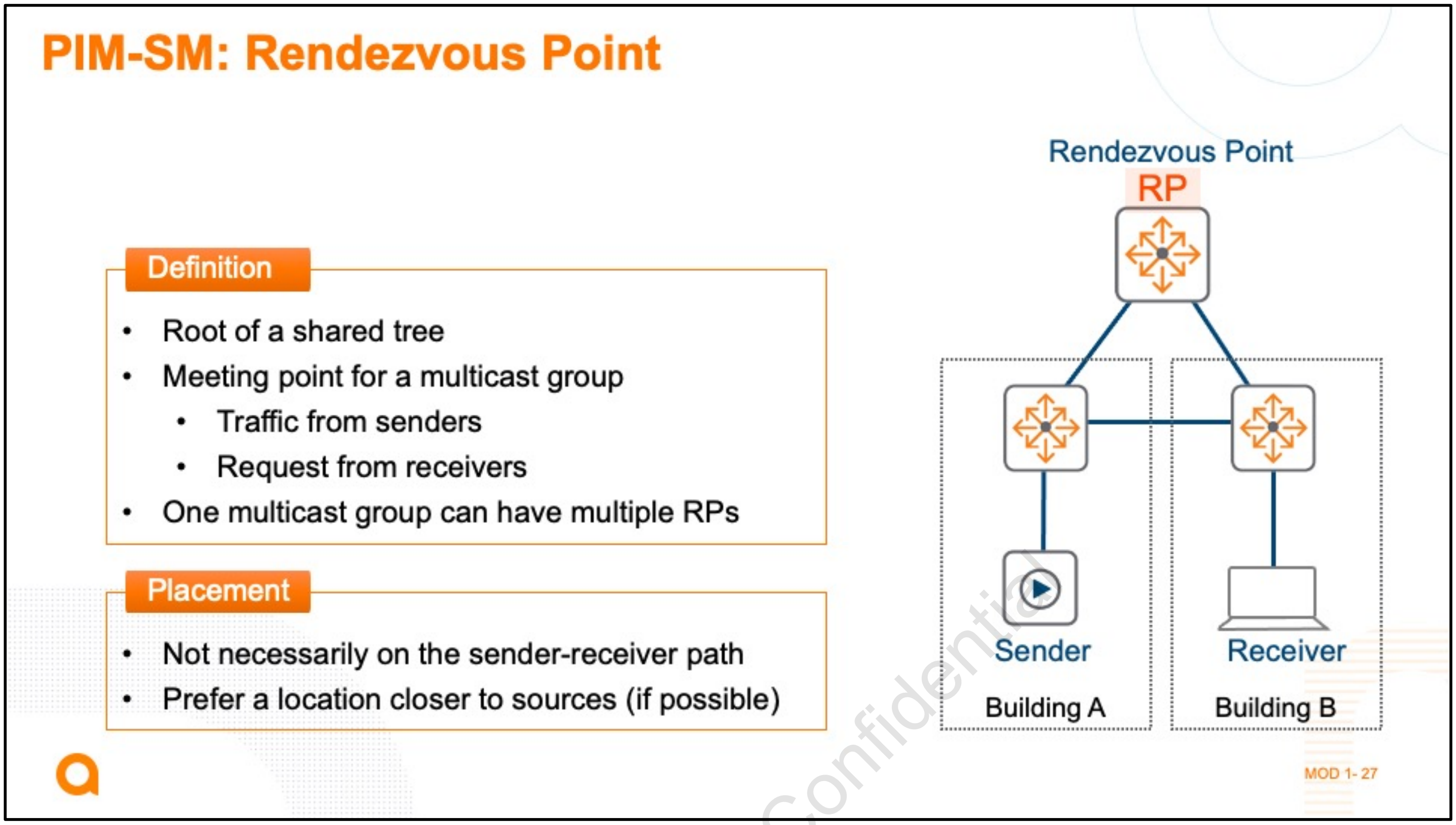
- Use priority value in Hello messages
  - Higher value wins, Default value = 1
  - If same priority, highest IP address wins
- When receiving an IGMP membership report
  - Sends PIM JOIN messages toward the root
- DR election on all interface types (p-to-p, broadcast)

A single network interface and IP address may connect to a multipoint or broadcast network, such as Ethernet. If the network is PIM-enabled, many neighbor adjacencies can occur, one for each PIM-enabled IP neighbor. If every PIM router on the segment were to attempt to manage the tree state, confusion and excess traffic could occur. Much like an OSPF DR, the PIM DR manages state information and PIM updates for all the routers on the segment.

The DR selection process occurs when all neighbors on a segment have replied to PIM hello messages. Each router on a segment selects one router to function as the DR. The DR router is chosen using the PIM DR priority parameter, where the highest priority router is selected as the DR. The DR priority is configurable and sent in the PIM hello message. If the DR priority value is not supplied by all routers, or the priorities match, the router with the highest IP address wins - the IP address of the interface sending the PIM message. By default, all interfaces have a PIM DR priority of 1. All routers on the segment should select the same DR router to avoid conflicts.

When the DR receives an IGMP membership report message from a receiver for a new group or source, the DR creates a tree to connect the receiver to the source. To do this, it sends a PIM join message out the interface toward the agreed-upon root of the multicast tree - the rendezvous point (RP). When the DR determines that the last host has left a group or source, it sends a PIM prune message to remove the path from the distribution tree.

From the RFC 4601 section 4.3, DR election occurs on all interface types - point-to-point or broadcast: ”Because the distinction between LANs and point-to-point interfaces can sometimes be blurred, and because routers may also have multicast host functionality, the PIM-SM specification makes no distinction between the two. Thus, DR election will happen on all interfaces, LAN or otherwise.”



PIM-SM: Rendezvous Point

Definition

- Root of a shared tree
- Meeting point for a multicast group
  - Traffic from senders
  - Request from receivers
- One multicast group can have multiple RPs

Placement

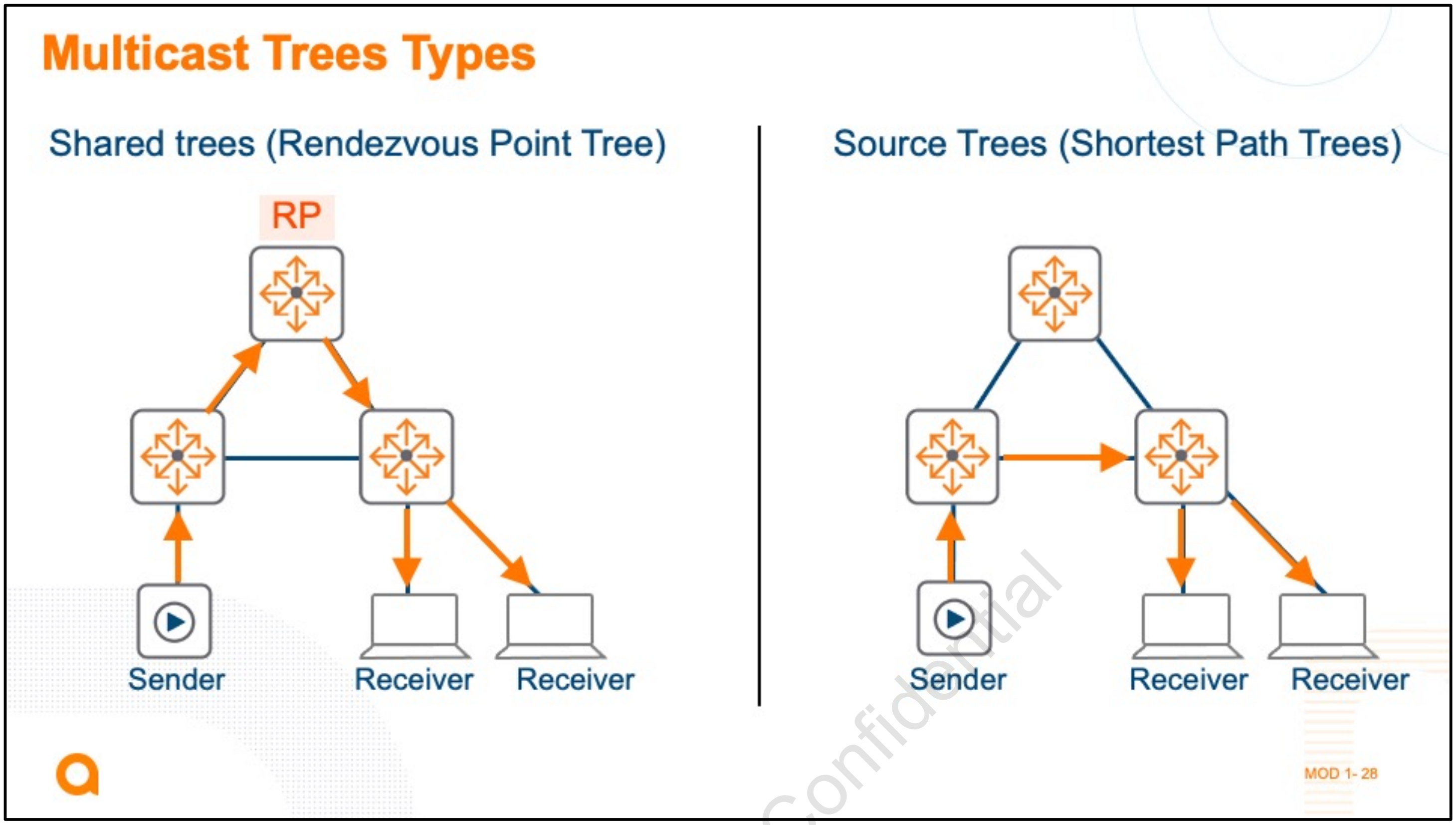
- Not necessarily on the sender-receiver path
- Prefer a location closer to sources (if possible)

MOD 1- 27

A Rendezvous Point (RP) is a router that has been configured to be used as the root of the non-source-specific distribution tree (Shared tree) for a multicast group. Join messages from receivers for a group are sent towards the RP, and data from senders is sent to the RP. Thus, receivers discover senders and start to receive traffic destined for the group. The RP router is not a packet source but simply a location from which to root and calculate a loop-free topology. The RP can also offload the processing of network joins and leaves during the initial setup of a multicast stream, saving network resources by concentrating them in a single location.

You should place the RP close to the source if possible. This is only applicable to a few sources that are of key importance to the business. Enterprise-wide deployments for multicast normally use RPs in the data center for the enterprise scope. The RP might not be in the path between sender and receivers. PIM-SM uses the RP to build a shared tree as an initial state, but it can move to an optimized shortest path tree. This means that traffic is not sent to RP all the time and so RP location has minimal impact on network performance.

Let’s look compare shared vs. shortest path trees.



Multicast Trees Types

Shared trees (Rendezvous Point Tree)

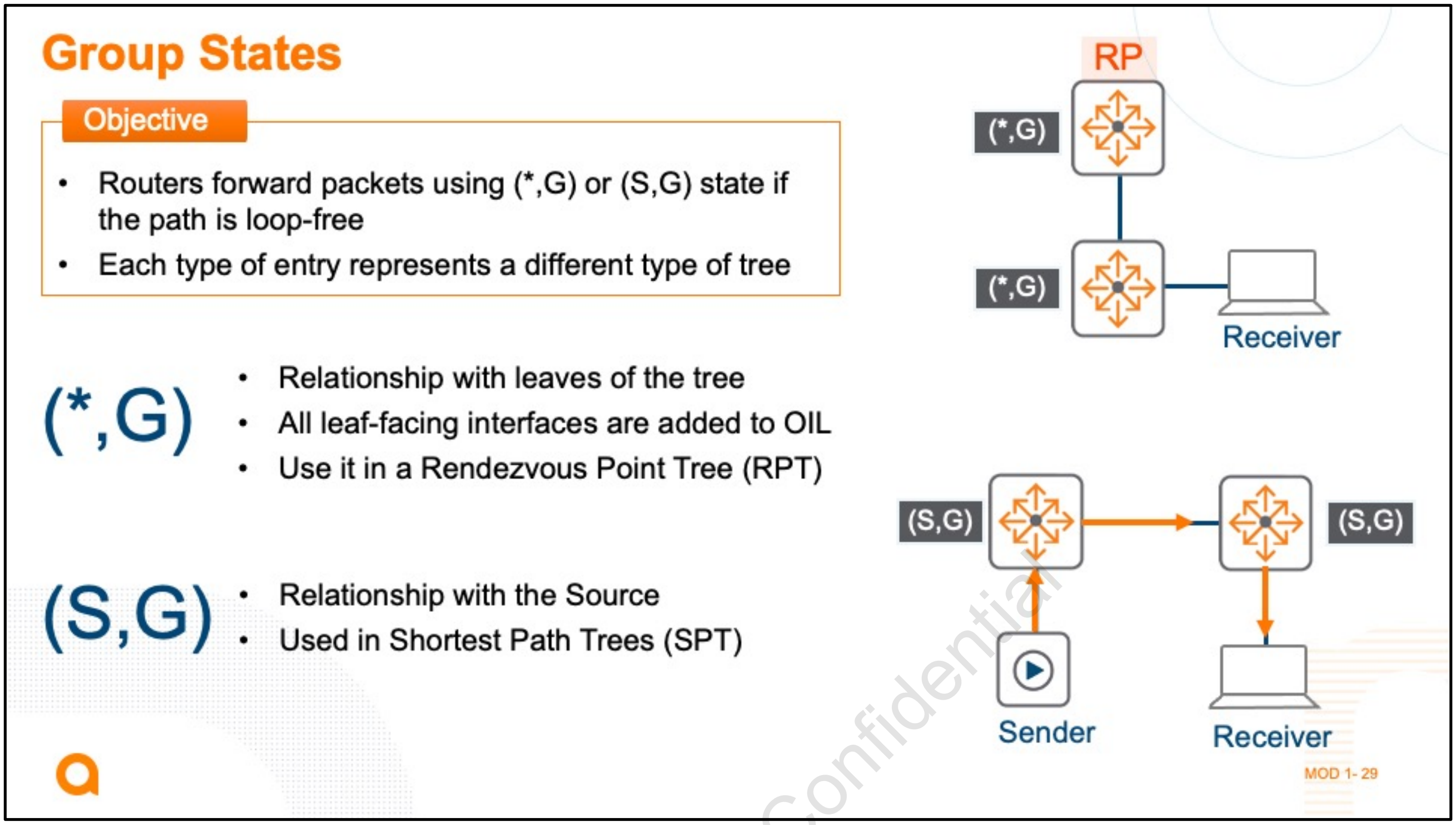
Source Trees (Shortest Path Trees)

MOD 1- 28

The two types of trees used in multicast networks to control the distribution of traffic are source trees and shared trees.

Source trees (Shortest Path Trees). This is the most common type of multicast tree. From a network perspective, the root of the source tree is the router closest to the source. The tree extends from the root router in a direct path toward all the group members. This creates a tree in which the path might look like the one shown in the figure. Each router in the network must calculate the tree independently based on the information it has received, whether by dynamic protocol updates (PIM), or by configuration. Using Reverse Path Forwarding (RPF), the router checks for loops and creates the Outgoing Interface List (OIL) from the best (sometimes called shortest) paths from the unicast routing table. For this reason, a source tree is also called a shortest path tree.

Shared trees (Rendezvous Point Tree - RPT). The shared tree uses a shared point in the network from which to build the distribution tree. This location is called a Rendezvous Point(RP) and is therefore the root of the tree. Each router still uses Reverse Path Forwarding (RPF) checks to keep the topology loop-free, but the RPF check is in the direction of the RP, not the source. Notice the RP placement has no relation to the IP multicast source. The RP may or may not be in the direct path (shortest path) between the group source and the group client(s). In fact, it can be placed outside of the path entirely. For this reason, a shared tree is also sometimes called a Rendezvous Point Tree (RPT). The advantage to using this type of tree is that all non-RP routers can conserve control- plane resources during stream establishment or during host maintenance. In addition, the network path is pre-determined and predictable, which leads to consistent network operations.



Group States

Objective

- Routers forward packets using (*,G) or (S,G) state if the path is loop-free
- Each type of entry represents a different type of tree

(*,G)

- Relationship with leaves of the tree
- All leaf-facing interfaces are added to OIL
- Use it in a Rendezvous Point Tree (RPT)

(S,G)

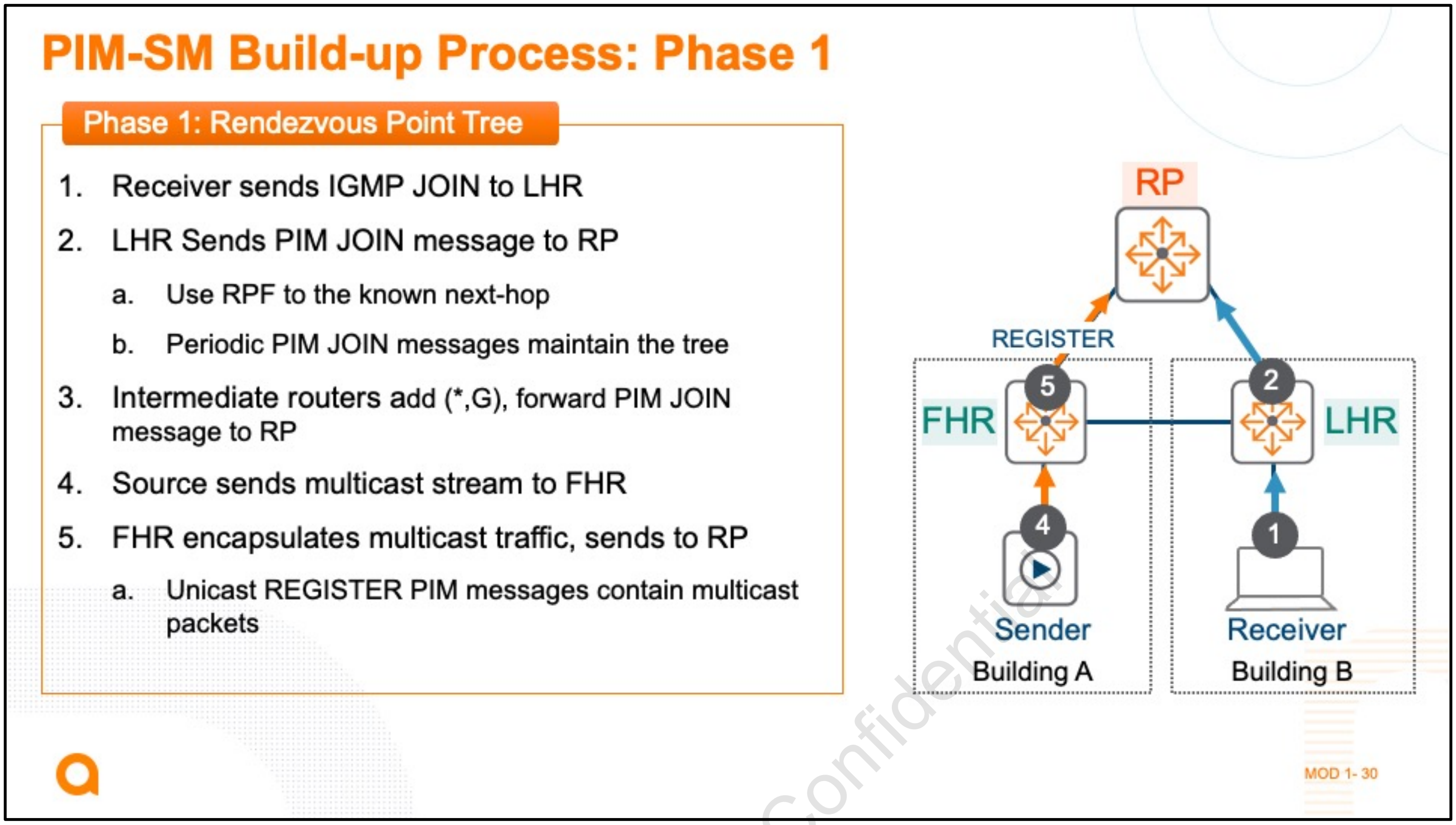
- Relationship with the Source
- Used in Shortest Path Trees (SPT)

Entries in the multicast routing table (mroute table) have the form (source,group).

A (*,G) entry in an mroute table represents a router's relationship to the leaves of a tree. All the leaf-facing interfaces are added to an outgoing interface list (OIL). This type of entry is used in a Rendezvous Point Tree (RPT) because routers have a preconfigured meeting point - the Rendezvous Point (RP). This means that routers need to only get one interested receiver for a particular multicast group to build the tree - the source is not really needed. The RP requests this information from the source.

The (S,G) entry in the mroute table represents the router's relationship to the source of the multicast stream. This type of entry is used in Shortest Path Trees. In this case not only the multicast group is need it but also the source IP address to

Routers can forward packets using either the (*,G) or the (S,G) entries, as long as the paths are determined to be loop-free.



PIM-SM Build-up Process: Phase 1

Phase 1: Rendezvous Point Tree

1. Receiver sends IGMP JOIN to LHR
2. LHR Sends PIM JOIN message to RP
   a. Use RPF to the known next-hop
   b. Periodic PIM JOIN messages maintain the tree
3. Intermediate routers add (*,G), forward PIM JOIN message to RP
4. Source sends multicast stream to FHR
5. FHR encapsulates multicast traffic, sends to RP
   a. Unicast REGISTER PIM messages contain multicast packets

MOD 1- 30

RFC 4601 defines three phases for the build-up process of PIM-SM.

Phase 1: Rendezvous Point Tree

Receiver endpoint sends an IGMP JOIN request asking for specific multicast group to the Last-Hop Router (LHR). This message includes only the Multicast group, therefore multicast state (*,G) is created on the FHR's multicast routing table. This action also adds the receiver-facing interface to the Outgoing Interface List.

LHR Sends a JOIN PIM message to the RP:

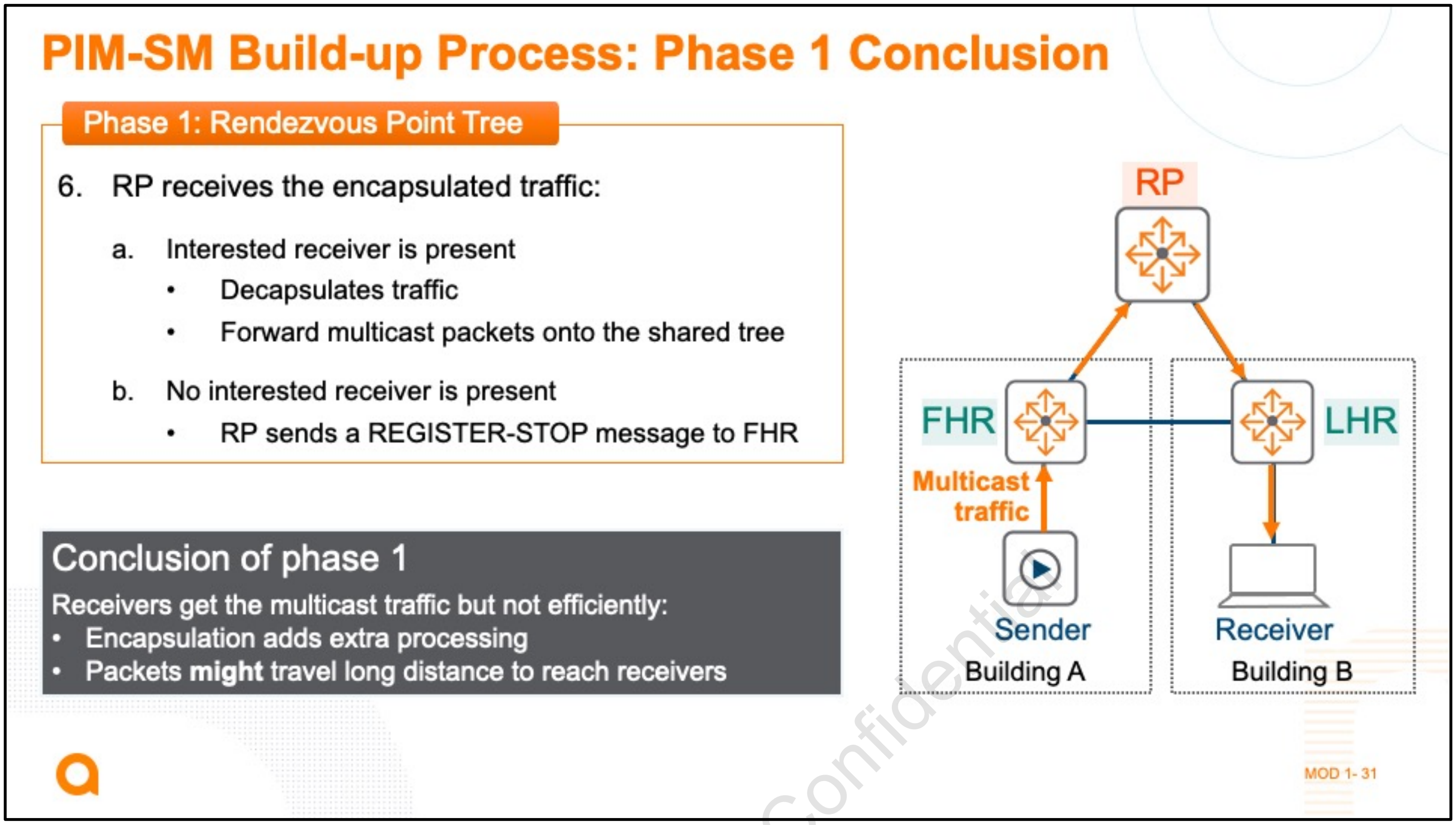
- LHR uses Reverse Path First (RPF) to look in the Unicast routing table and determine the next-hop to reach the RP.
- PIM Join messages are sent periodically to maintain the multicast state.

If intermediate routers exist between the LHR and RP, those devices add a (*,G) state on their multicast routing table and forward the PIM JOIN message to the RP. In the example this step is not needed.

Eventually the packet reaches the RP, which adds a (*,G) state in its multicast routing table and adds the receiver-facing interface in the OIL.

Source sends the multicast stream to the First Hop Router (FHR), which adds a state (S,G) to its multicast routing table.

The FHR must send the multicast traffic to the RP, but there is a problem. Intermediate routers (between the FHR and the RP path) and the RP do not have a (S,G) state in their multicast routing table. This means that routers will drop the multicast traffic coming from the FHR. To resolve this problem, the router closest to the source must encapsulate the multicast packet inside of a unicast packet with the IP address of the rendezvous point as the IP destination. In other words, the multicast stream must be tunneled from the router at the source to the RP. The encapsulation packets are known as PIM register packets.



PIM-SM Build-up Process: Phase 1 Conclusion

Phase 1: Rendezvous Point Tree

6. RP receives the encapsulated traffic:
   a. Interested receiver is present
      - Decapsulates traffic
      - Forward multicast packets onto the shared tree
   b. No interested receiver is present
      - RP sends a REGISTER-STOP message to FHR

Conclusion of phase 1

Receivers get the multicast traffic but not efficiently:

- Encapsulation adds extra processing
- Packets **might** travel long distance to reach receivers

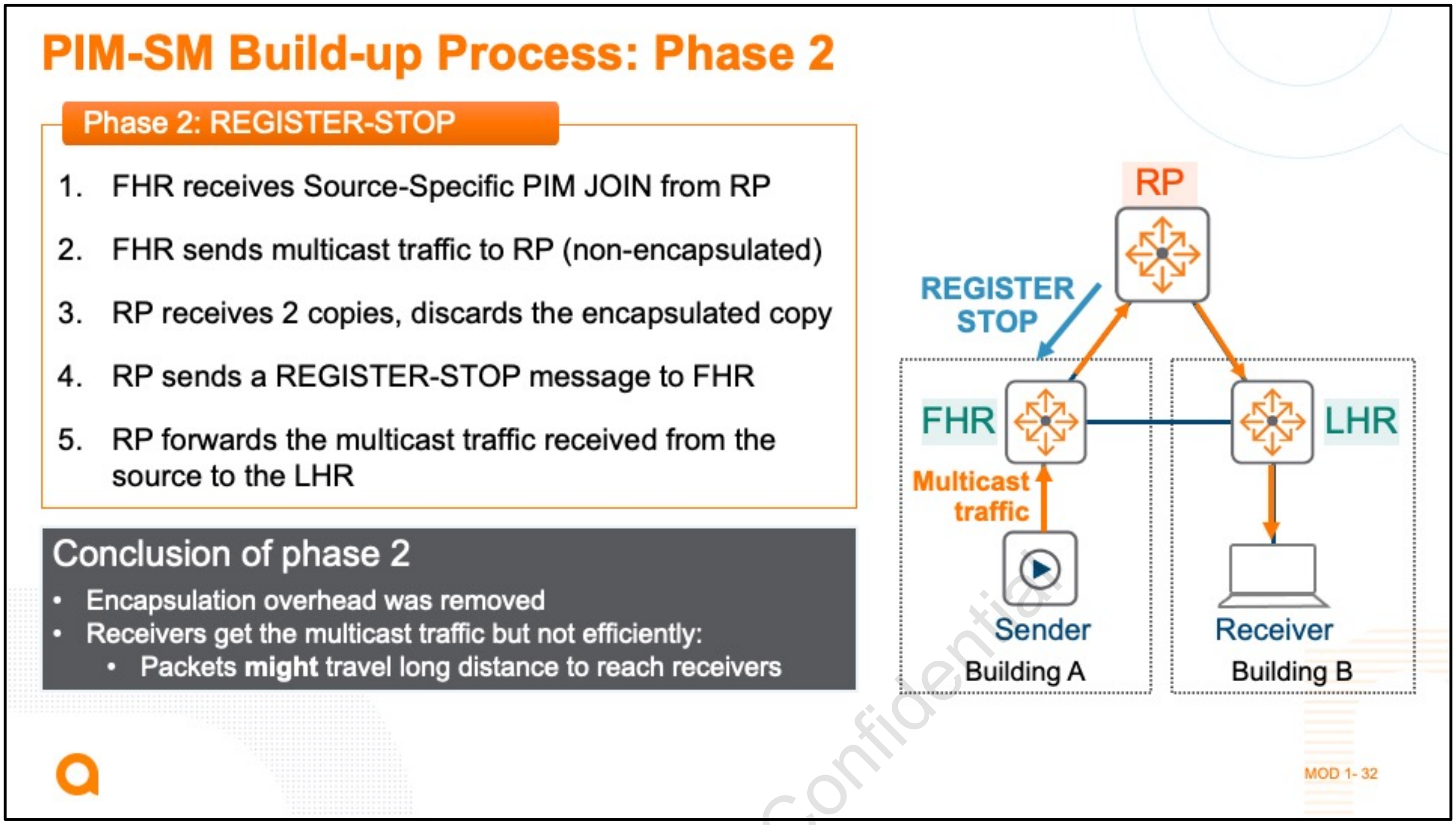
MOD 1- 31

Phase 1: Rendezvous Point Tree

6. The RP receives the PIM REGISTER message (encapsulated traffic) and checks for interested receivers present (*,G) entry in its multicast routing table.

- If there is a receiver present, then it decapsulates the traffic and forward the multicast packets onto the shared tree.
- If there is no reciver present, then the RP sends a REGISTER-STOP message to the FHR. Each router (between RP and source) maintains a register suppression timer that is initiated by the REGISTER STOP message. Upon expiration of this timer, the FHR again sends a register packet to the RP to verify if there are any active receivers for that group

At the end of phase one, multicast traffic flows encapsulated to the RP, and then natively over the RP tree to the multicast receivers. During this time, the flow is not efficient.



PIM-SM Build-up Process: Phase 2

Phase 2: REGISTER-STOP

1. FHR receives Source-Specific PIM JOIN from RP
2. FHR sends multicast traffic to RP (non-encapsulated)
3. RP receives 2 copies, discards the encapsulated copy
4. RP sends a REGISTER-STOP message to FHR
5. RP forwards the multicast traffic received from the source to the LHR

Conclusion of phase 2

- Encapsulation overhead was removed
- Receivers get the multicast traffic but not efficiently:
  - Packets **might** travel long distance to reach receivers

MOD 1- 32

Phase 2: REGISTER-STOP

The FHR receives a source-specific PIM JOIN message from the RP. This adds a (S,G) entry to the FHR multicast routing table and to the intermediate routers in the path. This state permits all devices to process multicast information that is received from the source, this means that the encapsulation process is no longer required.

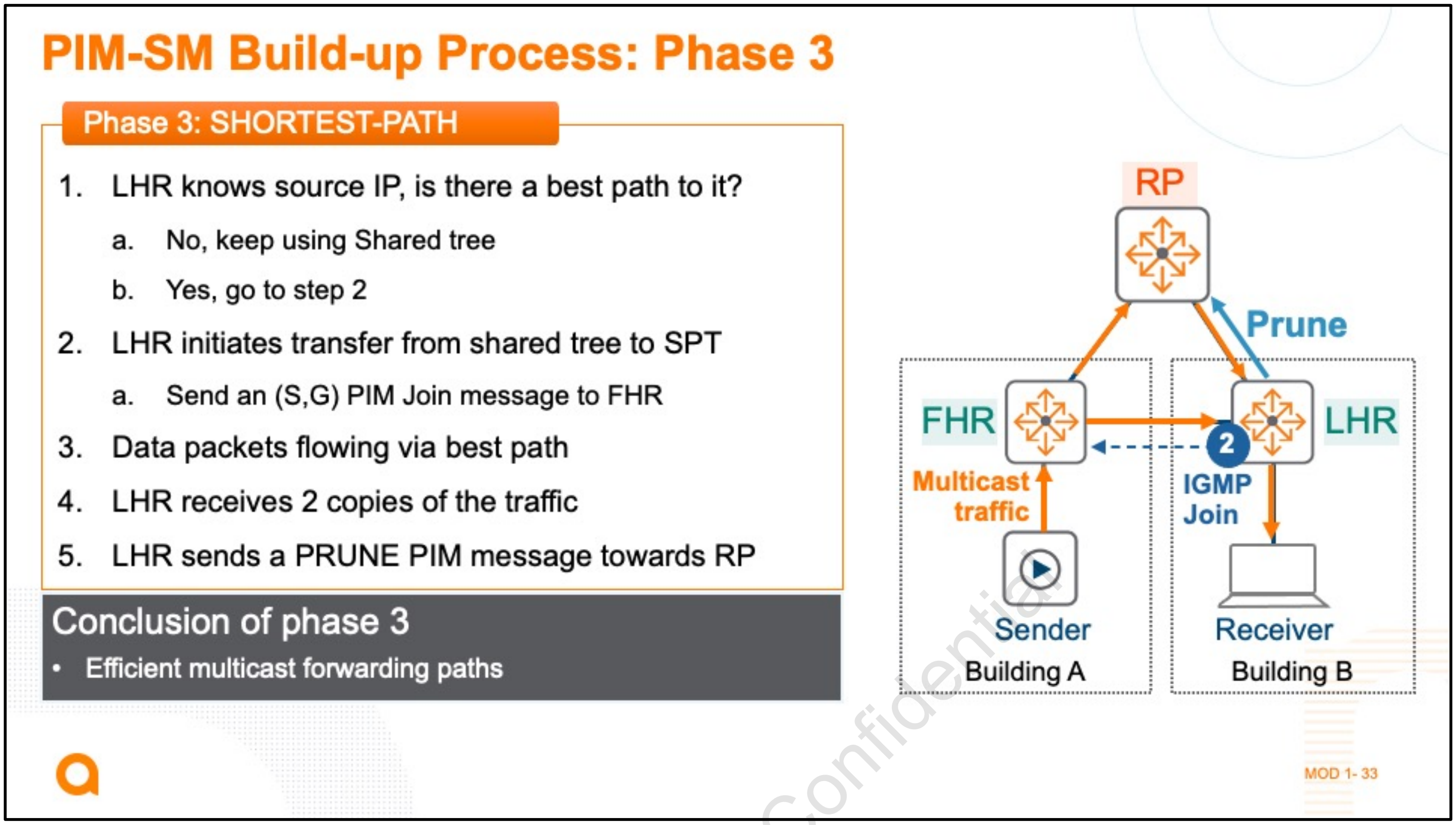
The FHR sends the multicast traffic natively (non-encapsulated) to the RP. All intermediate devices in the path can route the traffic.

At this moment, the RP receives two copies of the same traffic, one encapsulated in the REGISTER message and other natively received. The RP discards the encapsulated copy.

The RP sends a PIM Register-Stop messsage to FHR.

RP receives the multicast traffic from the source with no encapsulation and forwards this traffic to the LHR, this action creates a new state (S,G) in the multicast routing table on all routers between the RP and the LHR.

At the end of phase 2, encapsulation overhead has been removed but the multicast traffic might travel long distance to reach receivers.



PIM-SM Build-up Process: Phase 3

Phase 3: SHORTEST-PATH

1. LHR knows source IP, is there a best path to it?
   a. No, keep using Shared tree
   b. Yes, go to step 2
2. LHR initiates transfer from shared tree to SPT
   a. Send an (S,G) PIM Join message to FHR
3. Data packets flowing via best path
4. LHR receives 2 copies of the traffic
5. LHR sends a PRUNE PIM message towards RP

Conclusion of phase 3
- Efficient multicast forwarding paths

MOD 1- 33

Phase 3. Shortest Path

Now LHR knows the multicast traffic source IP address, and asks itself if there a best path to that source, using the RPF process. If there is no a new best path available, then the LHR keeps using the shared tree path. But if there is a new best path, then a path optimization process must be performed – Step 2.

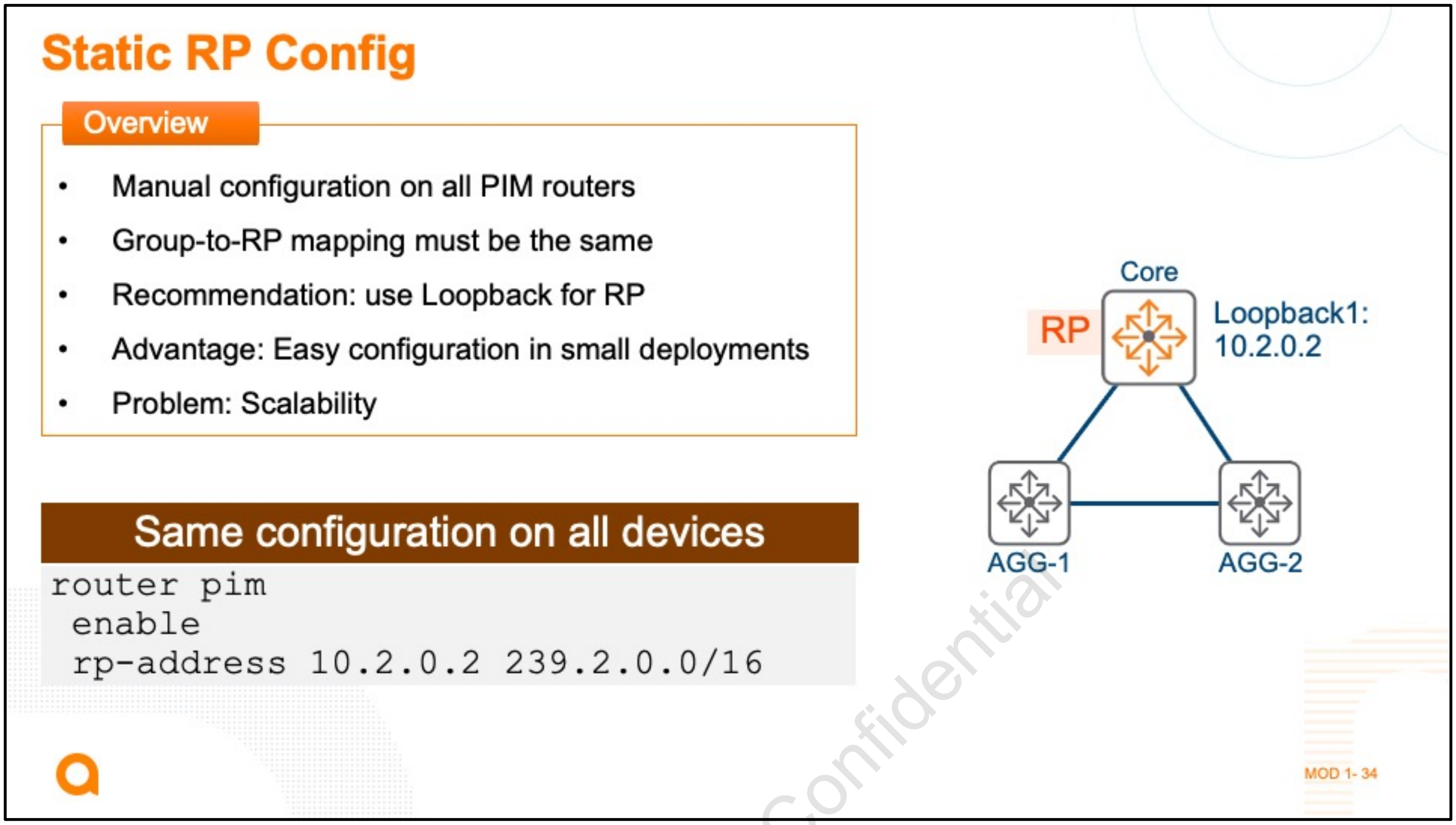
The LHR initiates a transfer from the shared tree to the Shortest Path Tree. The LHR sends a PIM Join message (S,G) to the FHR.

When the FHR receives the PIM message it starts forwarding multicast packets to the LHR via the best path.

LHR receives two copies of the traffic, one from the RP using the shared tree and other from the Source using the SPT.

LHR sends a PIM Prune message towards RP.

At the end of phase 3, receivers get the multicast traffic efficiently.



Static RP Config

Overview

- Manual configuration on all PIM routers
- Group-to-RP mapping must be the same
- Recommendation: use Loopback for RP
- Advantage: Easy configuration in small deployments
- Problem: Scalability

Same configuration on all devices

```
router pim
 enable
 rp-address 10.2.0.2 239.2.0.0/16
```

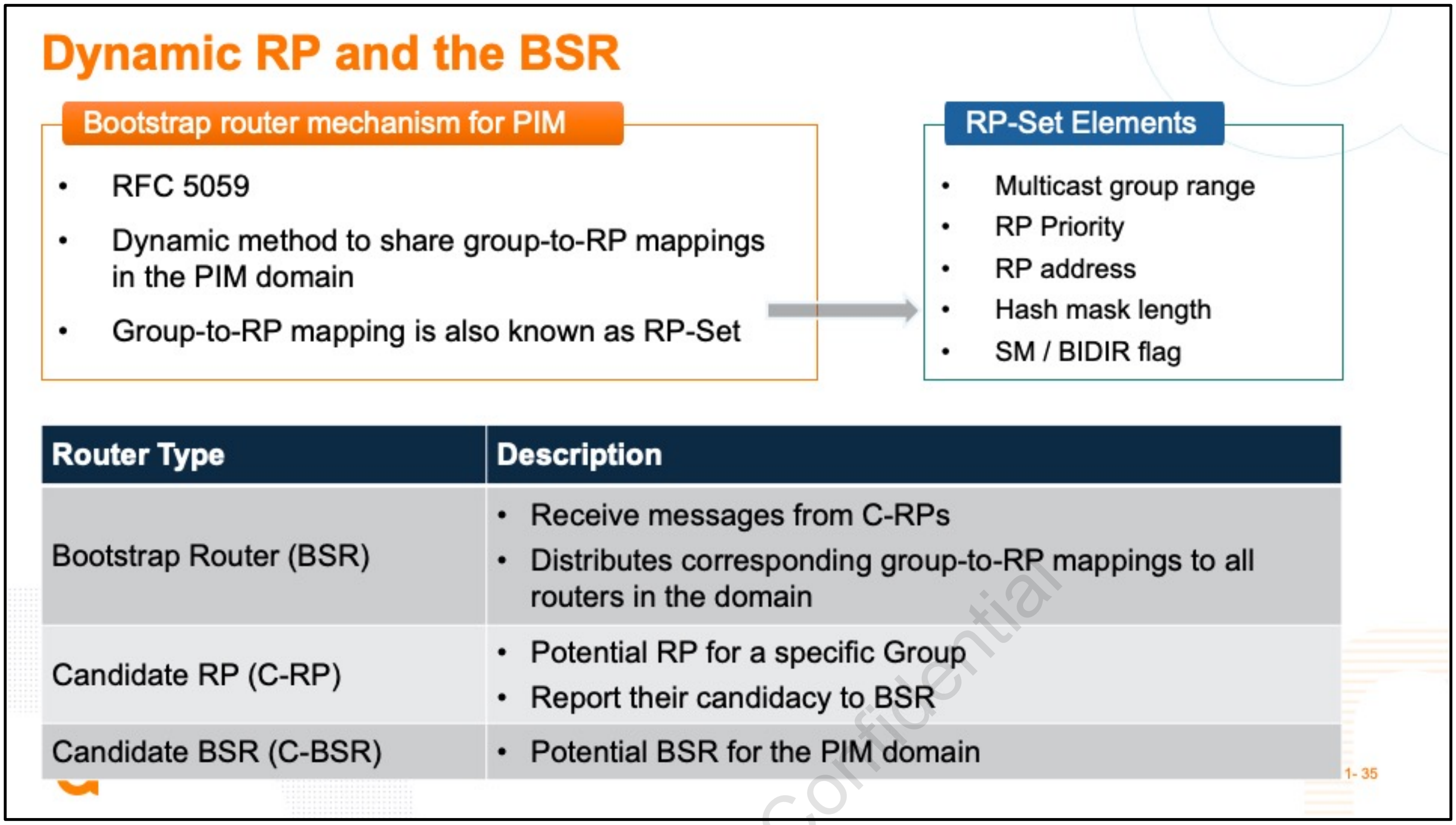
MOD 1- 34

Enabling PIM sparse-mode using a static RP is perhaps the simplest way to configure a multicast domain. You must manually configure this – telling every device the IP address of the RP – including the desired RP. You are mapping a device to function as the RP for a particular group address. This configuration must be defined identically on every router. Otherwise, the domain will be incomplete.

The configuration shown in the figure statically configures the router as the RP for a specified multicast group or range of multicast groups. This must be configured on all PIM-SM routers in the domain. If group address is not specified, it applies to all IPv4 multicast addresses (224.0.0.0 - 239.255.255.255). PIM-SM supports a maximum of 8 static RPs per VRF.

Why is it important to configure the main network loopback interface with sparse-mode PIM? This is a recommended practice for any multicast overlay. This ensures that the router can fully participate in the multicast domain, even if errors occur on leaf-facing interfaces. It also allows the loopback interfaces to be used as a RP addresses or mapping agents in dynamic RP propagation, making them more reliable.

This static configuration option can be suitable in small deployments with few multicast group mappings. However, for large deployments with many overlapping domains or many multicast applications and many rendezvous points, a consistent group mapping through static commands may become extremely difficult to manage. You may need a more scalable solution than static configuration – the Bootstrap Router mechanism.



Dynamic RP and the BSR

Bootstrap router mechanism for PIM

- RFC 5059
- Dynamic method to share group-to-RP mappings in the PIM domain
- Group-to-RP mapping is also known as RP-Set

RP-Set Elements

- Multicast group range
- RP Priority
- RP address
- Hash mask length
- SM / BIDIR flag

Router Type	Description
Bootstrap Router (BSR)	• Receive messages from C-RPs • Distributes corresponding group-to-RP mappings to all routers in the domain
Candidate RP (C-RP)	• Potential RP for a specific Group • Report their candidacy to BSR
Candidate BSR (C-BSR)	• Potential BSR for the PIM domain

1- 35

The BootStrap Router (BSR) mechanism for PIM is defined in RFC 5059.

BSR is a RP high-availability mechanism that provides active/standby functionality and automatic downstream RP information propagation. This mechanism dynamically estabishes a method to share group-to-RP mappings (also known as RP-Set) to all PIM routers in the domain. This mechanism avoids wrong mappings that can cause unexpected problems.

With BSR, you configure some routers within a PIM domain to be potential RPs – so-called Candidate-RPs (C-RPs). A subset of the C-RPs will eventually be used as the actual RPs for the domain. You also configure some PIM routers in the domain to be candidate bootstrap routers, or Candidate-BSRs (C-BSRs). One C-BSRs is elected to be the BSR for the domain. All PIM routers in the domain learns the result of this election through Bootstrap messages. The C-RPs then report their candidacy to the elected BSR, which chooses a subset of these C-RPs and distributes corresponding group-to-RP mappings to all the routers in the domain through Bootstrap messages.

RP-Set elements shared by Bootstrap are:

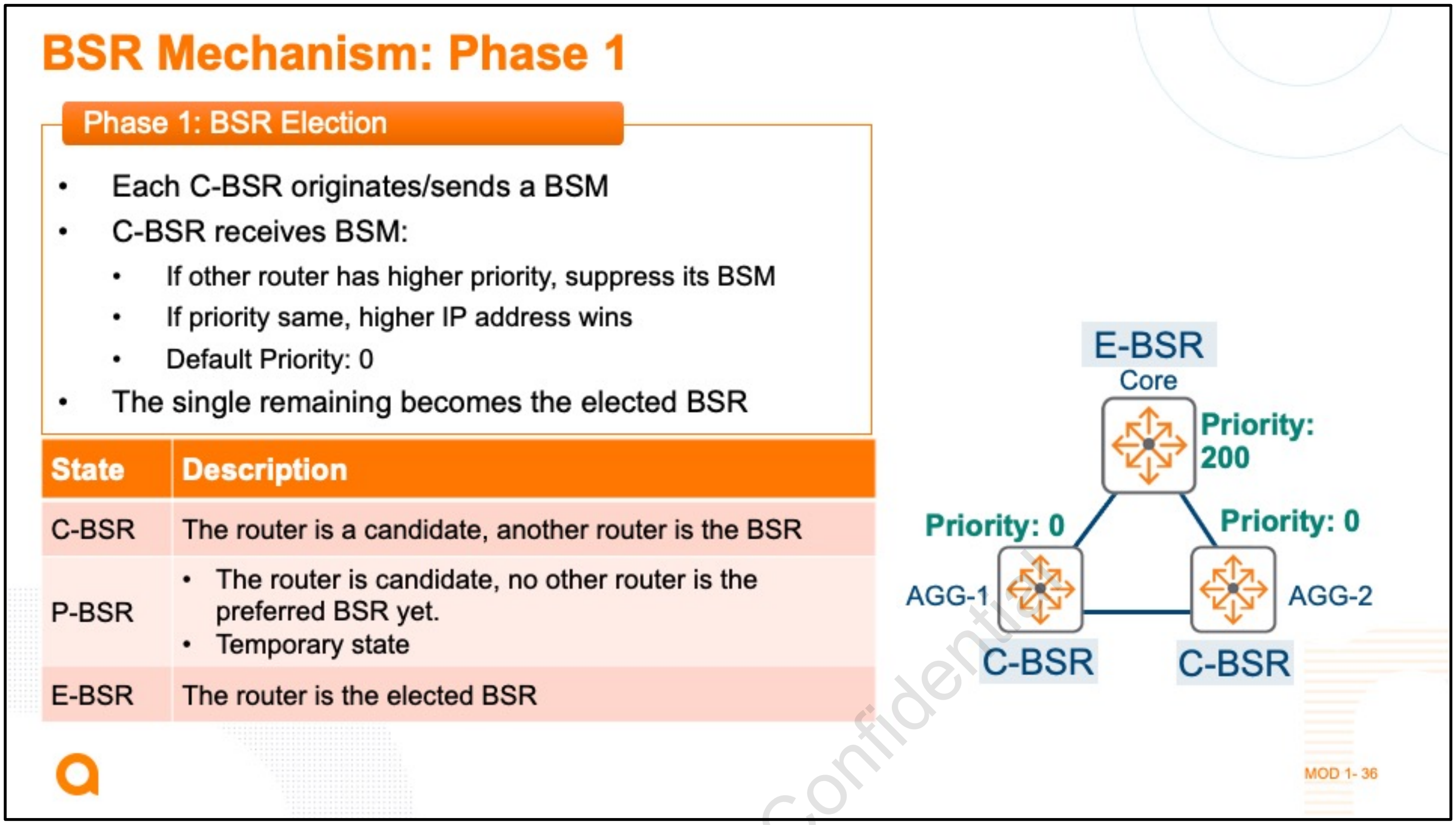
Multicast group range, expressed as an address and prefix length

RP Priority

RP address

Hash mask length

SM / BIDIR flag



BSR Mechanism: Phase 1

Phase 1: BSR Election

- Each C-BSR originates/sends a BSM
- C-BSR receives BSM:
  - If other router has higher priority, suppress its BSM
  - If priority same, higher IP address wins
  - Default Priority: 0
- The single remaining becomes the elected BSR

State	Description
C-BSR	The router is a candidate, another router is the BSR
P-BSR	• The router is candidate, no other router is the preferred BSR yet. • Temporary state
E-BSR	The router is the elected BSR

MOD 1- 36

The bootstrap process has four phases as defined in RFC 5059.

Phase 1 - BSR Election

The function of the BSR is to collect and broadcast the RP set to all routers in the domain. Each Candidate-BSR originates BootStrap Messages (BSMs) announcing their capacity to become a BSR. Every BSM contains a BSR Priority field. Routers within the domain flood the BSMs throughout the domain. A Candidate or C-BSR that hears about a higher-priority C-BSR than itself stops sending BSMs for some period of time. The single remaining C-BSR becomes the elected BSR, and its BSMs inform all the other routers in the domain that it is the elected BSR. If the priority value is the same, the router with the highest IP address is elected.

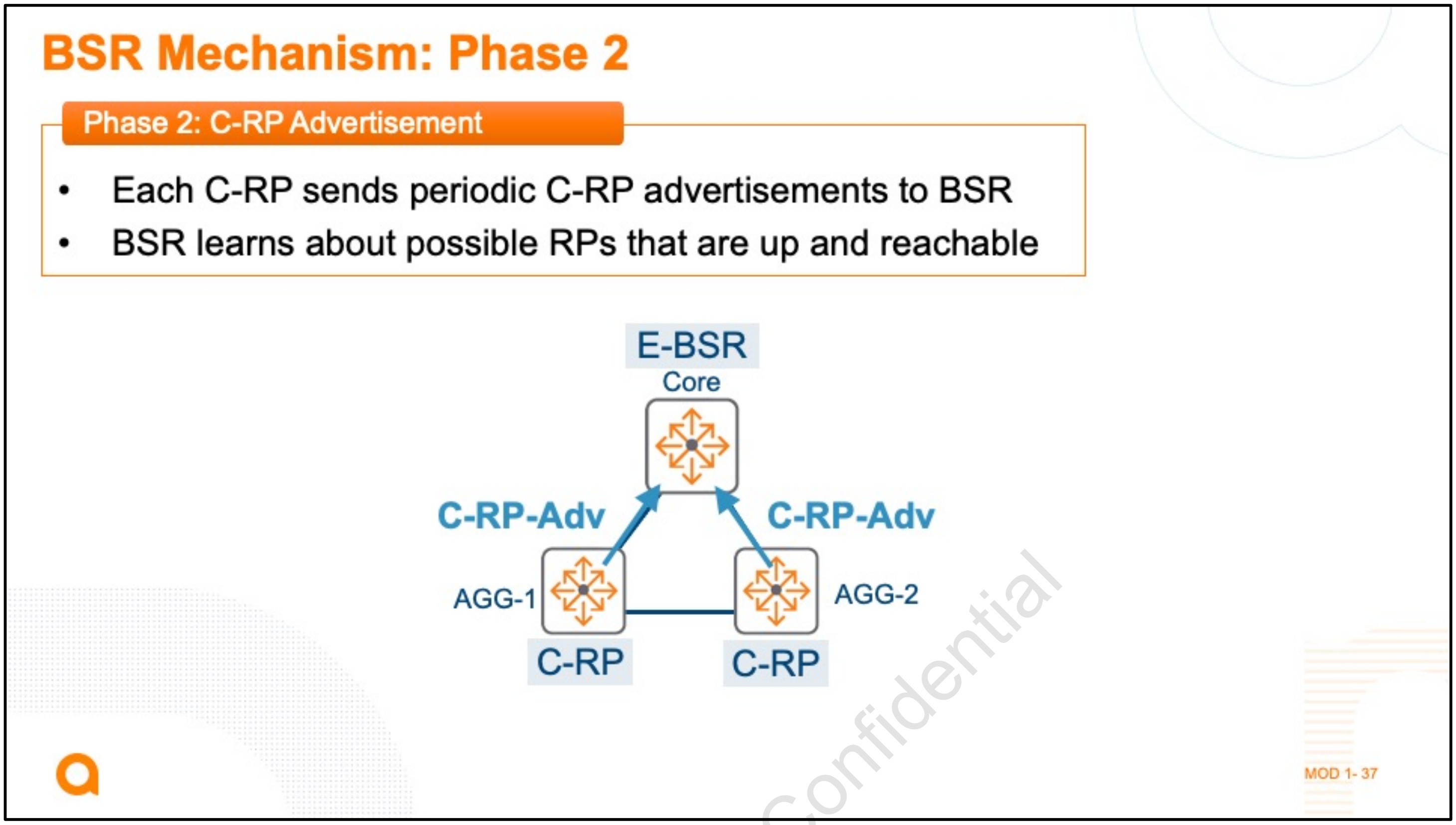
There are three states during the BSR election:

Candidate-BSR (C-BSR). The router is a candidate to be the BSR for the scope zone, but currently another router is the preferred BSR.

Pending-BSR (P-BSR) The router is a candidate to be the BSR for the scope zone. Currently, no other router is the preferred BSR, but this router is not yet the elected BSR. This is a temporary state that prevents rapid thrashing of the choice of BSR during BSR election.

Elected-BSR (E-BSR) The router is the elected BSR for the scope zone and it must perform all the BSR functions.

Note. A key part of the election mechanism is that we associate a weight with each BSR. The weight of a BSR is defined to be the concatenation in fixed-precision unsigned arithmetic of the BSR Priority field from the Bootstrap message and the IP address of the BSR from the Bootstrap message (with the BSR Priority taking the most-significant bits and the IP address taking the least-significant bits).



BSR Mechanism: Phase 2

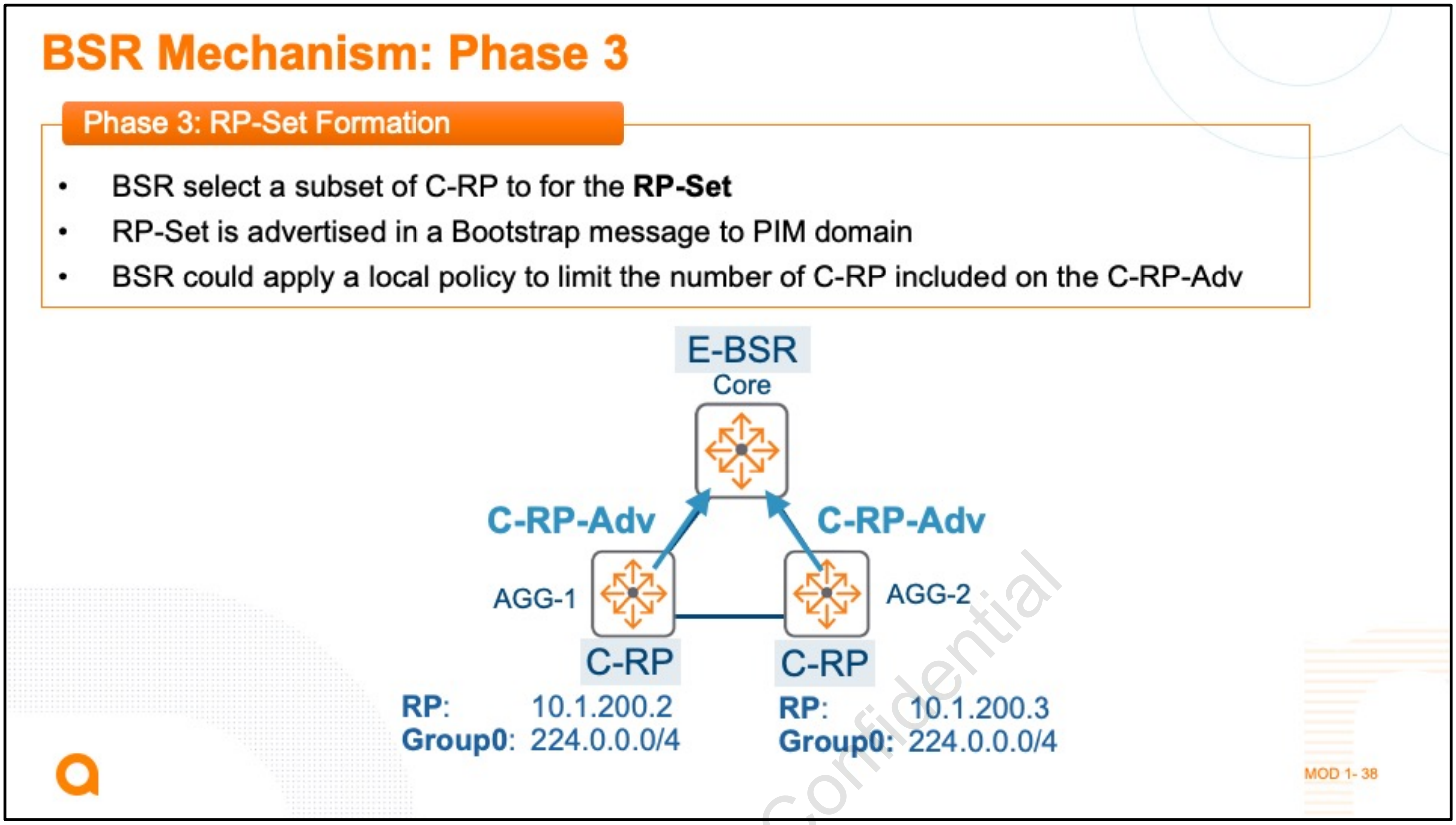
Phase 2: C-RP Advertisement

- Each C-RP sends periodic C-RP advertisements to BSR
- BSR learns about possible RPs that are up and reachable

MOD 1- 37

Phase 2 – C-RP Advertisment

Each Candidate-RP within a domain sends periodic Candidate-RP-Advertisement (C-RP-Adv) messages to the elected BSR. A C-RP-Adv message includes the priority of the advertising C-RP, as well as a list of group ranges for which the candidacy is advertised. In this way, the BSR learns about possible RPs that are currently up and reachable.



BSR Mechanism: Phase 3

Phase 3: RP-Set Formation

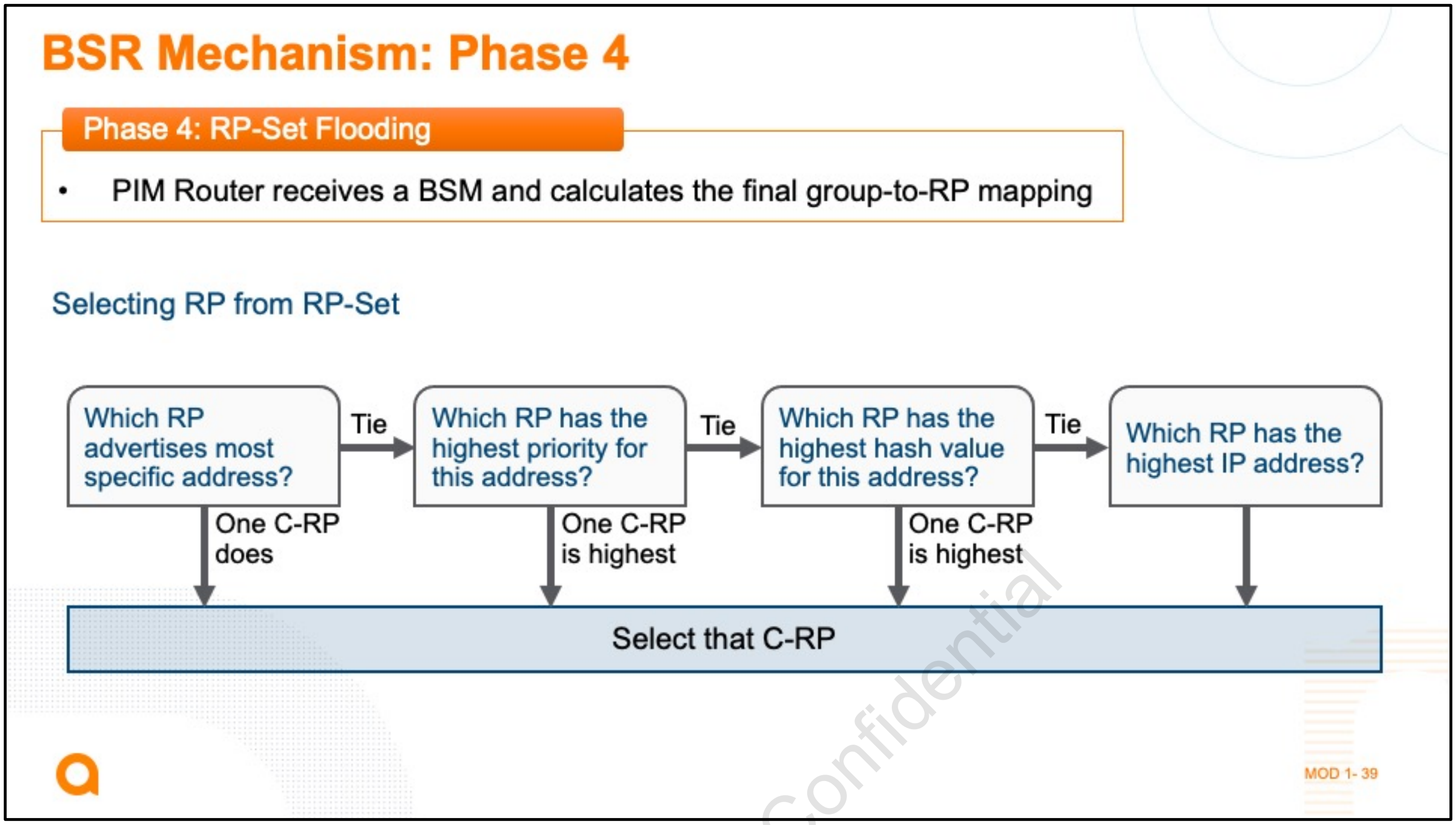
- BSR select a subset of C-RP to for the **RP-Set**
- RP-Set is advertised in a Bootstrap message to PIM domain
- BSR could apply a local policy to limit the number of C-RP included on the C-RP-Adv

MOD 1- 38

Phase 3 – RP-Set Formation

To form the RP-Set, the BSR selects a subset of the C-RPs from which it has received C-RP-Adv messages. In general, it should do this in such a way that the RP-Set is neither so large that all the routers in the domain cannot be informed about it, nor so small that the load is overly concentrated on some RPs. It should also attempt to produce an RP-Set that does not change frequently.

Note: For security reasons, there should be a way to restrict which IP addresses the BSR accepts C-RP-Adv messages from – such as an access list. For use of scoped BSR, it may also be useful to specify which group ranges should be accepted.



BSR Mechanism: Phase 4

Phase 4: RP-Set Flooding

- PIM Router receives a BSM and calculates the final group-to-RP mapping

Selecting RP from RP-Set

Which RP advertises most specific address?

Tie

Which RP has the highest priority for this address?

Tie

Which RP has the highest hash value for this address?

Tie

Which RP has the highest IP address?

One C-RP does

One C-RP is highest

One C-RP is highest

Select that C-RP

MOD 1- 39

Phase 4 – RP-Set Flooding

When a PIM router receives a Bootstrap message, it adds the group-to-RP mappings contained therein to its pool of mappings obtained from other sources (e.g., static configuration). It calculates the final mappings of group addresses to RP addresses using the following criteria:

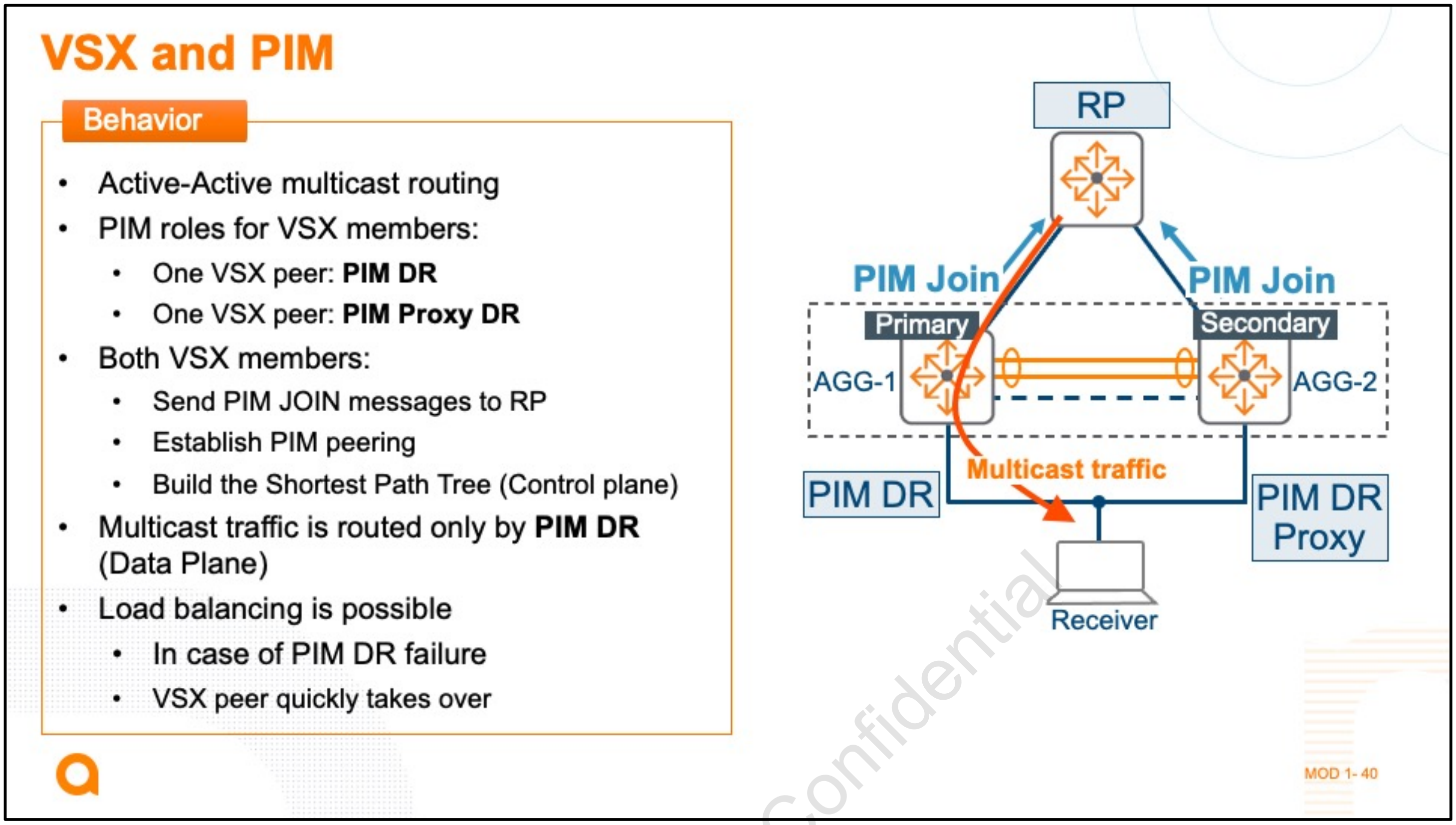
Select the RP that advertises the most specific range of multicast addresses that includes the desired address.

If more than one C-RP advertises a range of the same specificity, select the C-RP configured with the highest priority.

If more than one C-RP is configured with the highest priority, a hash function generates a series of mask length values that are individually assigned to the set of eligible C-RPs. If the hash function matches a single RP candidate to a longer mask length than the other candidates, that candidate is selected to support the group.

If the hash function matches the longest mask length to multiple RP candidates, the C-RP that has the highest IP address is selected to support the group.

In future Bootstrap messages, the BSR includes the RP-Set information. Bootstrap messages are flooded through the domain, which ensures that the RP-Set rapidly reaches all the routers in the domain. BSMs are originated periodically to ensure consistency after failure restoration.



VSX and PIM

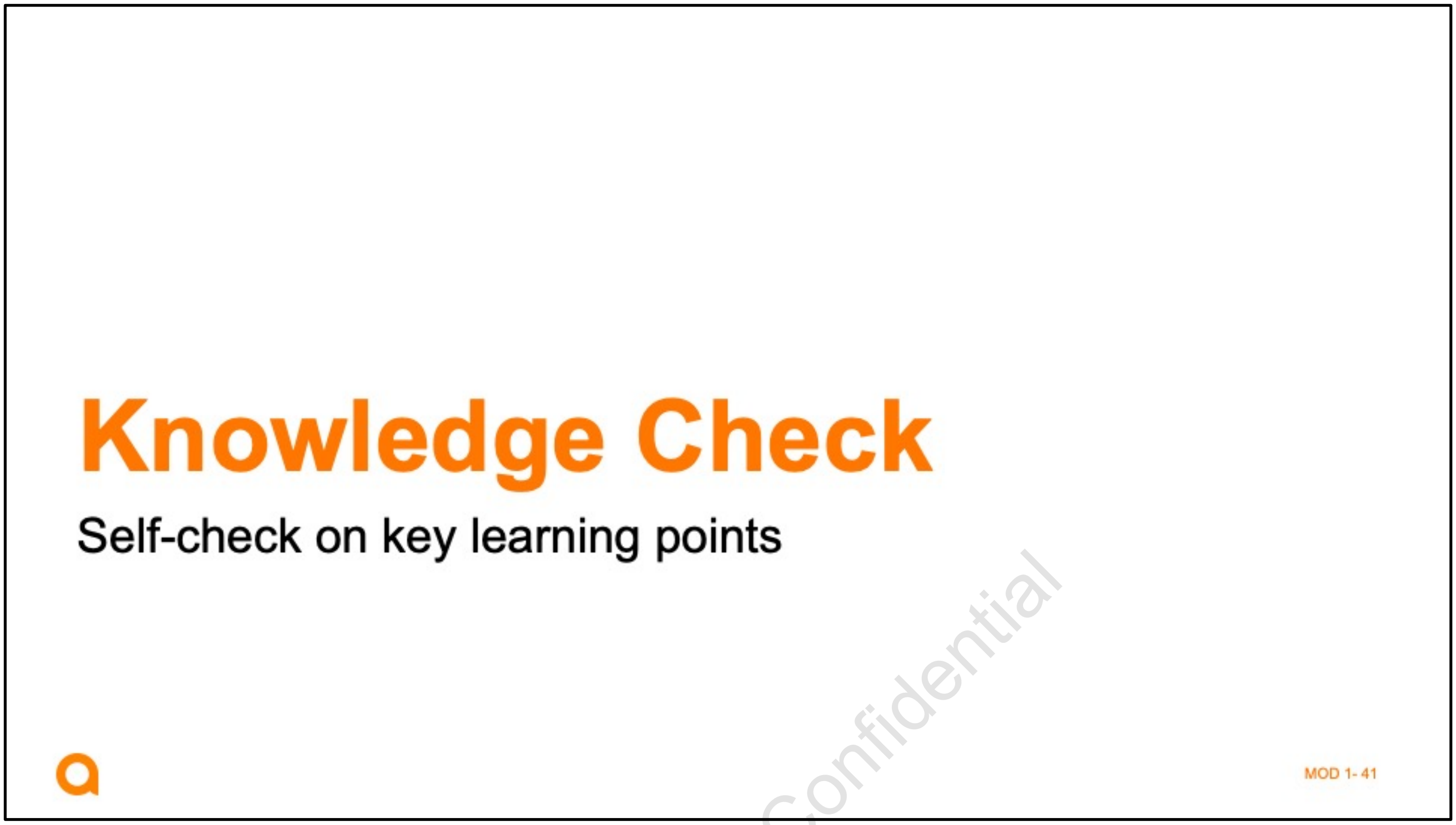
Behavior

- Active-Active multicast routing
- PIM roles for VSX members:
  - One VSX peer: **PIM DR**
  - One VSX peer: **PIM Proxy DR**
- Both VSX members:
  - Send PIM JOIN messages to RP
  - Establish PIM peering
  - Build the Shortest Path Tree (Control plane)
- Multicast traffic is routed only by **PIM DR** (Data Plane)
- Load balancing is possible
  - In case of PIM DR failure
  - VSX peer quickly takes over

AOS-CX supports dual designated router (DR) functionality for VSX. One VSX member is elected as PIM DR for each IP subnet, using the typical PIM election process. The other VSX member becomes a proxy DR.

Both VSX peers have the same Control Plane information. This means that both members will be able to establish PIM neighborships, send PIM Join messages to the RP and Build a Shortest Path Tree (SPT). However, multicast traffic (data plane) is only routed from the VSX peer that acts as the PIM DR. The mechanism to have a pre-established Control Plane on both VSX peers permits the VSX cluster to achieve a fast fail over in case the PIM DR fails.

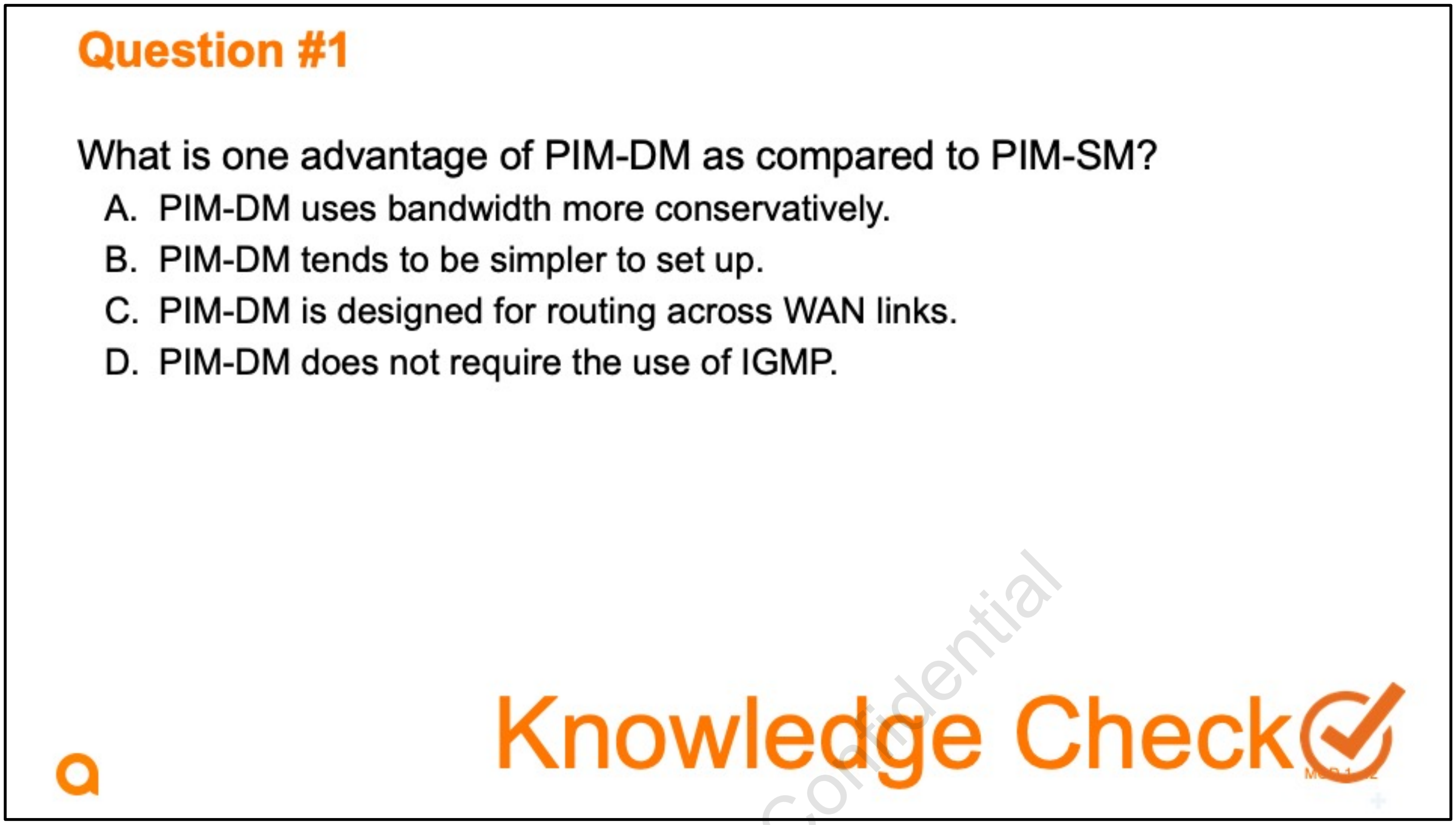
You can implement load balancing so that for each VSX pair, one switch can be the DR for a group of VLANs and the other switch the DR for the other group of the VLANs.



Knowledge Check

Self-check on key learning points

MOD 1- 41



Question #1

What is one advantage of PIM-DM as compared to PIM-SM?

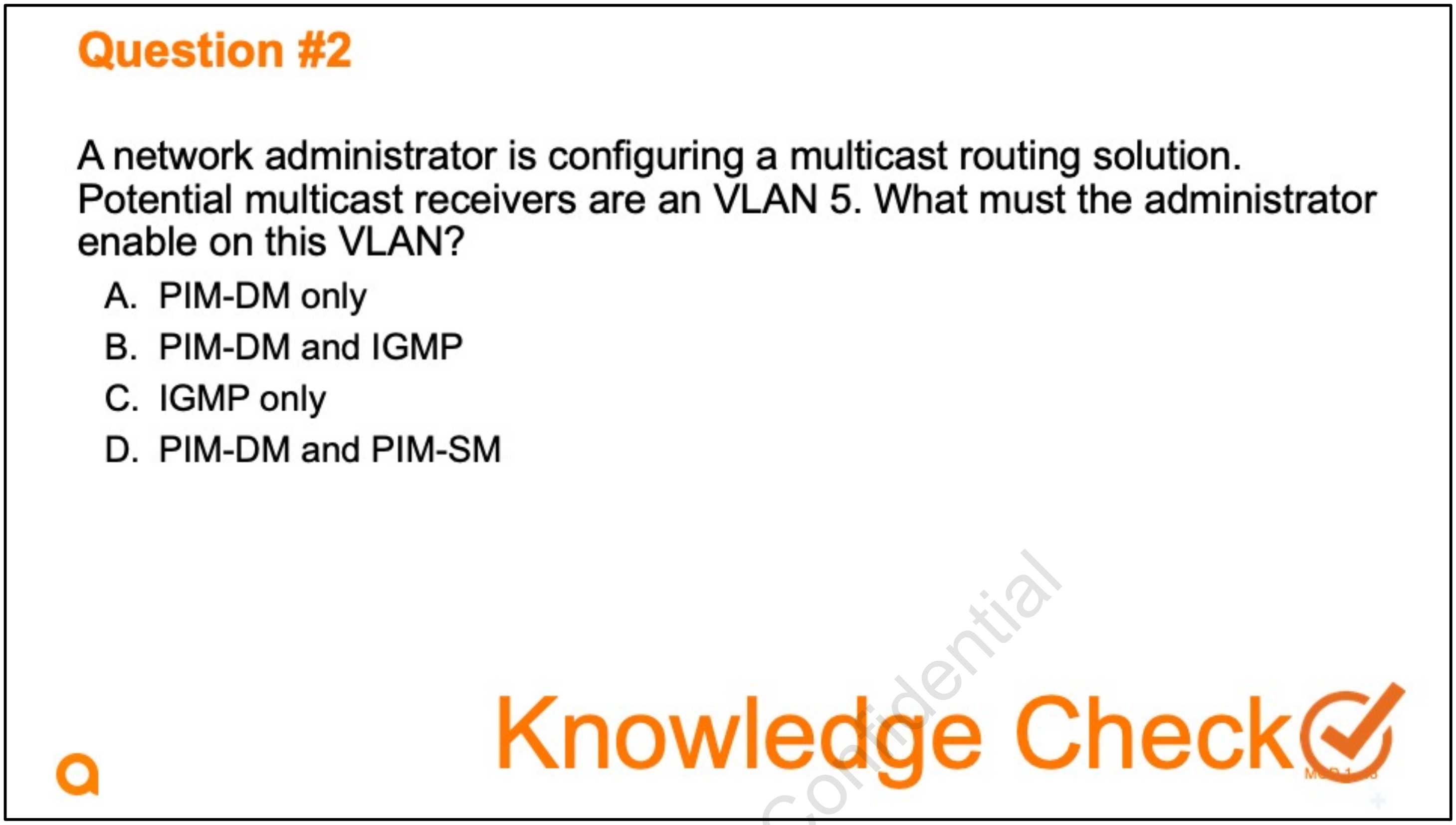
A. PIM-DM uses bandwidth more conservatively.

B. PIM-DM tends to be simpler to set up.

C. PIM-DM is designed for routing across WAN links.

D. PIM-DM does not require the use of IGMP.

Knowledge Check



Question #2

A network administrator is configuring a multicast routing solution. Potential multicast receivers are an VLAN 5. What must the administrator enable on this VLAN?

A. PIM-DM only

B. PIM-DM and IGMP

C. IGMP only

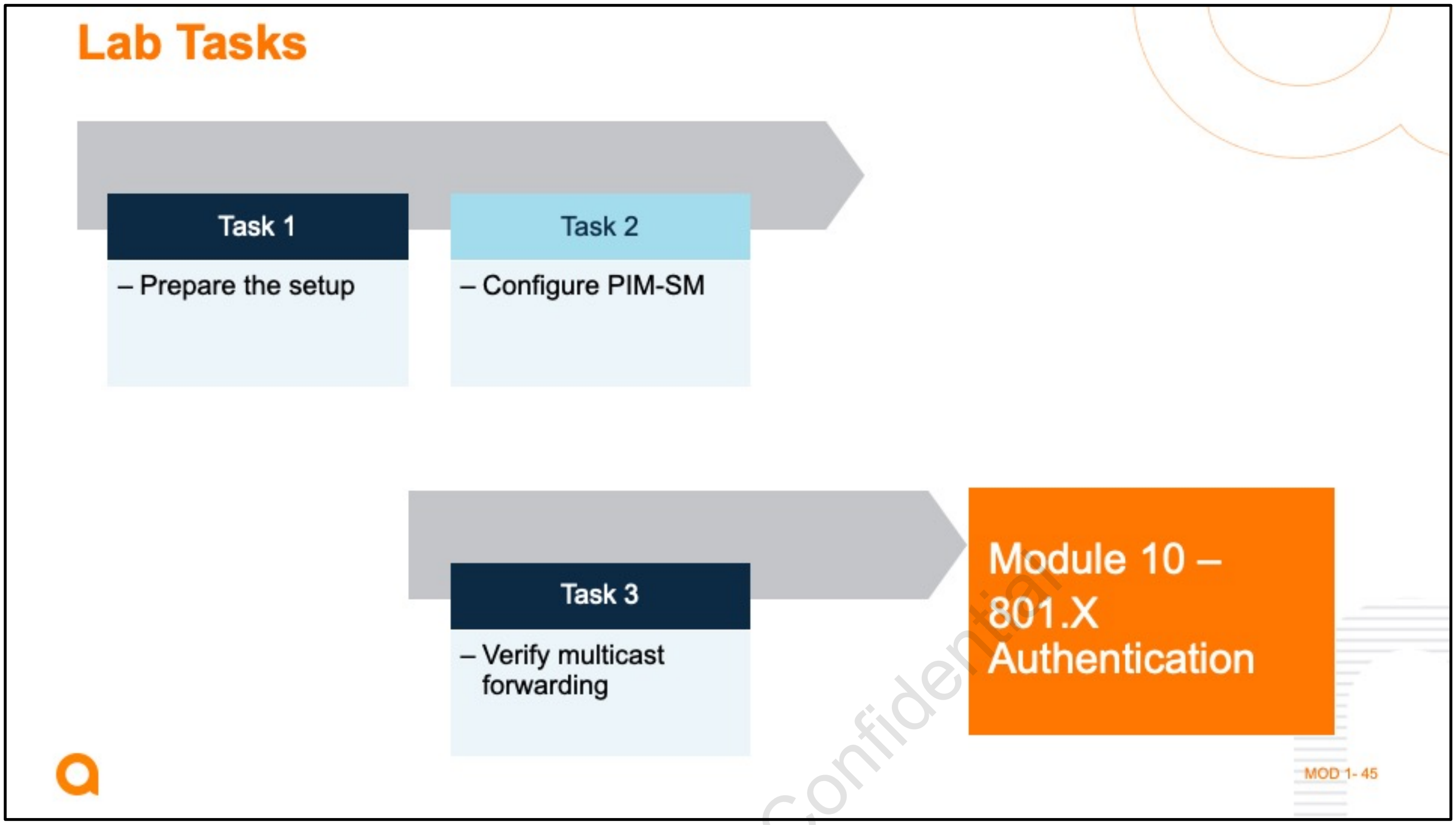
D. PIM-DM and PIM-SM

Knowledge Check

Lab Activity

Lab 6.3

Time for a lab.



Lab Tasks

Task 1

– Prepare the setup

Task 2

– Configure PIM-SM

Task 3

– Verify multicast forwarding

Module 10 – 801.X Authentication

MOD 1- 45

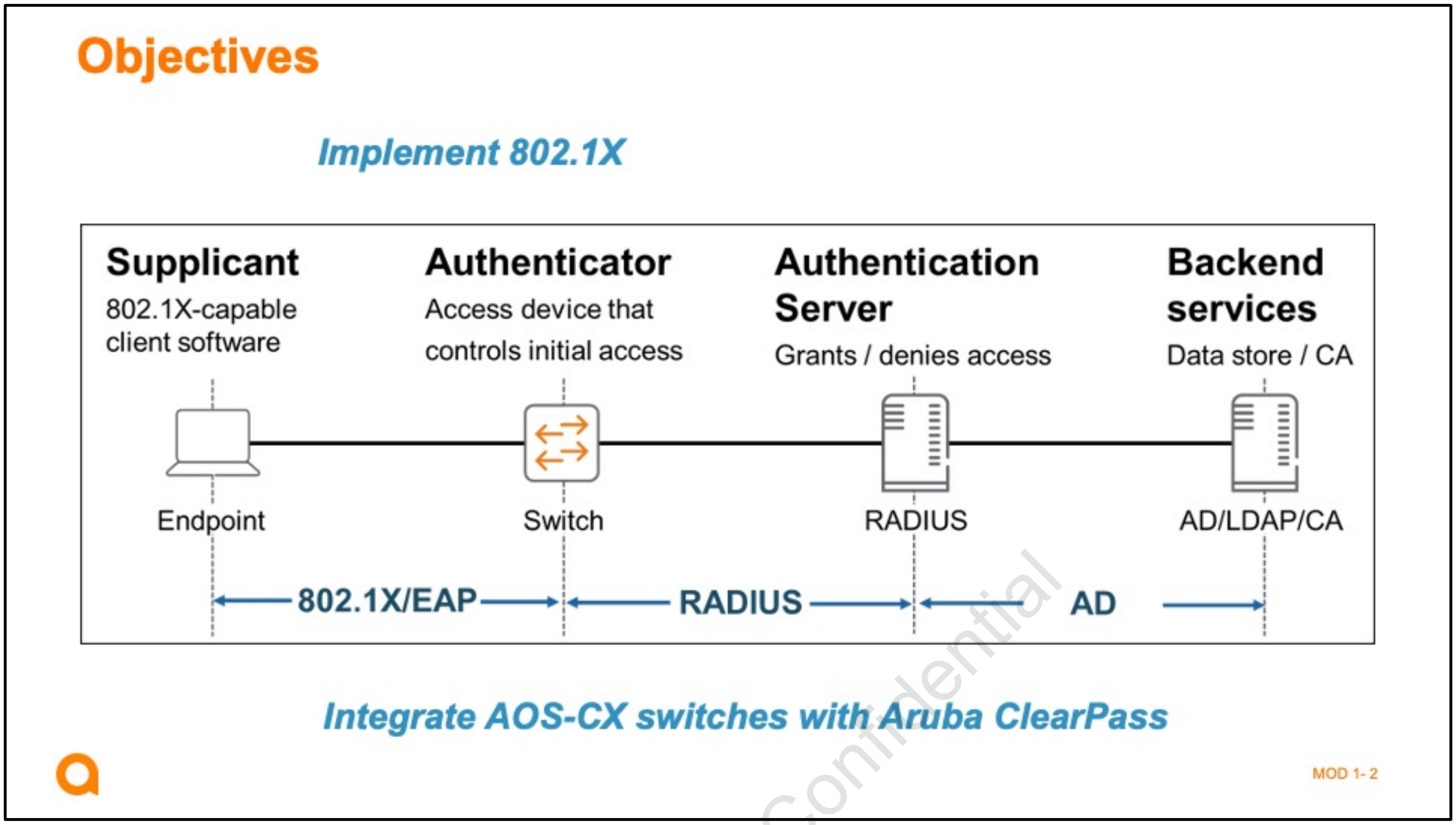
The figure provides a brief review of lab tasks. Please see your lab guide for details. When you are ready, please continue with Module 10 – 802.1X Authentication.

Rev # 20.21

802.1X Authentication

Implementing ArubaOS-CX Switching

You need endpoint security, so here’s the 802.1X Authentication module.



Objectives

Implement 802.1X

Supplicant
802.1X-capable client software

Authenticator
Access device that controls initial access

Authentication Server
Grants / denies access

Backend services
Data store / CA

Integrate AOS-CX switches with Aruba ClearPass

MOD 1- 2

By the time that you have completed this module, you should be able to:

Implement 802.1X on AOS-CX switch ports

Integrate AOS-CX switches with Aruba ClearPass. Why? dynamic VLAN assignments, ACLs, QoS priorities, and rate limits

Overview

- User Authentication
- Basic 802.1X
- Configuring 802.1X
- RADIUS Attributes
- User Roles
- Lab Activity

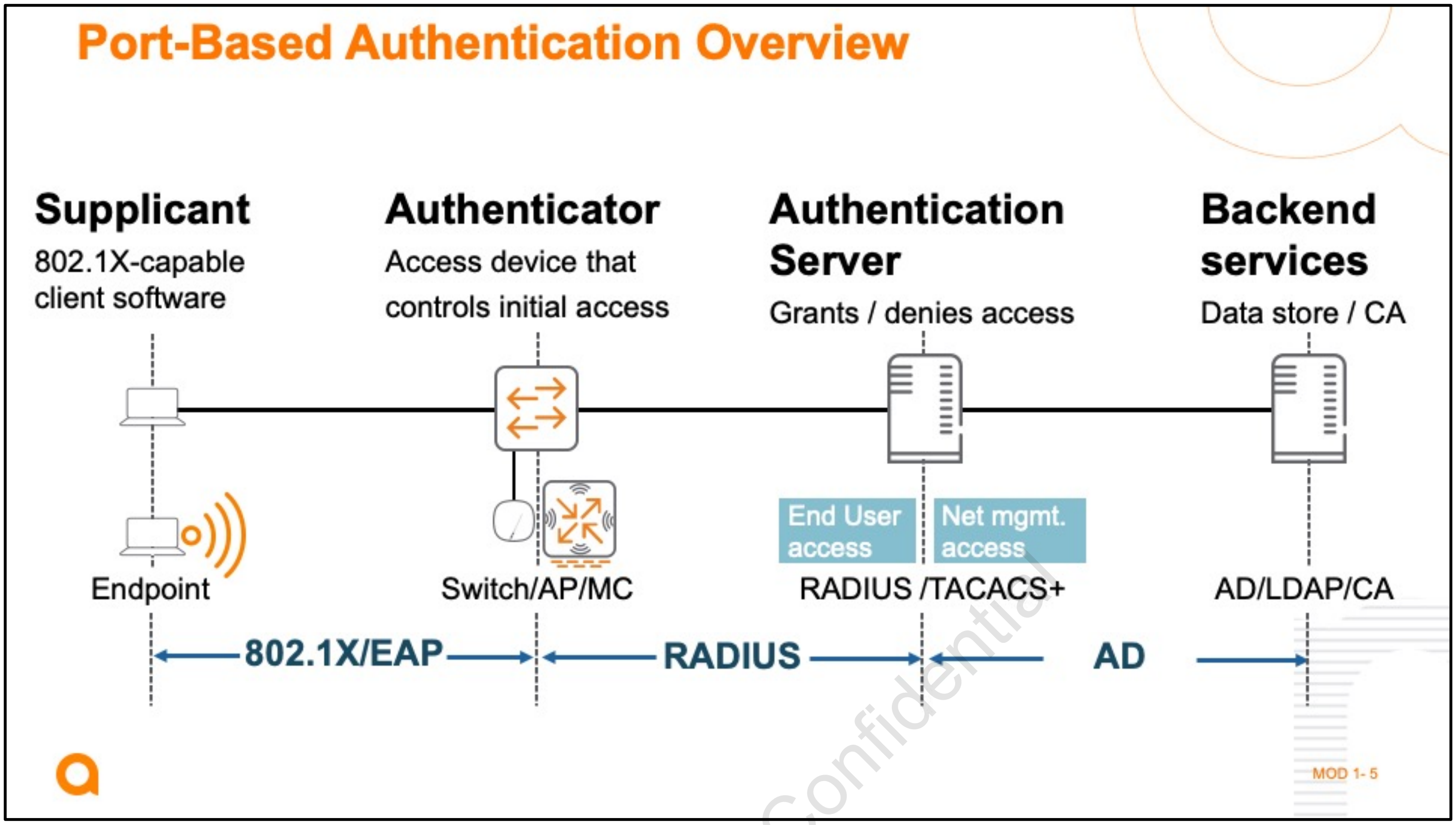
MOD 1- 3

This module introduces you to port-based authentication and teaches you how to implement 802.1X authentication technology on AOS-CX switches.

User Authentication

aruba
a Hewlett Packard Enterprise company

We begin with user authentication.



Port-Based Authentication Overview

Supplicant
802.1X-capable client software

Authenticator
Access device that controls initial access

Authentication Server
Grants / denies access

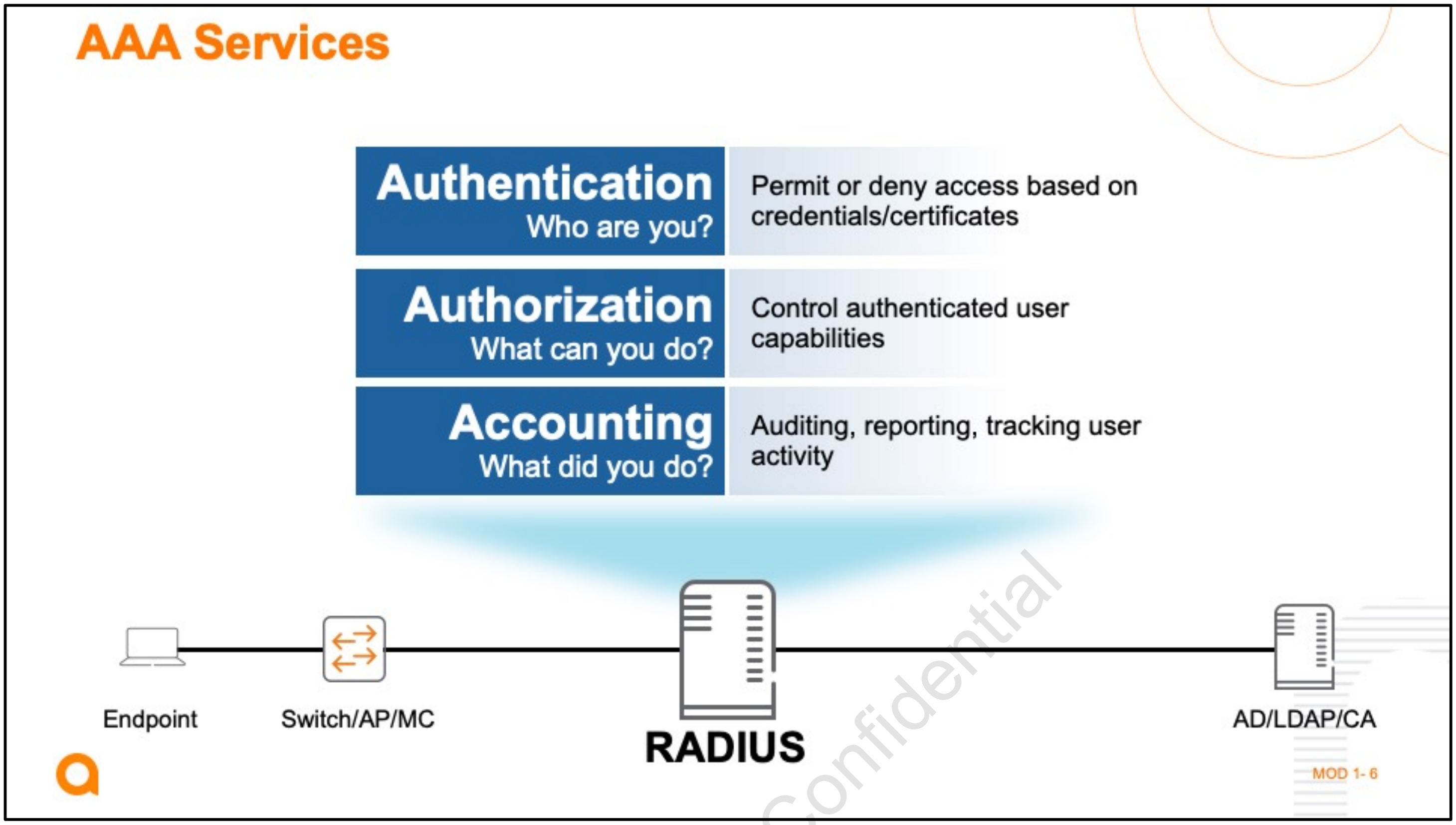
Backend services
Data store / CA

Let’s look at a typical port-based authentication scenario. You have endpoints that must connect to an access switch or AP to gain appropriate network access. For tight security, these endpoints often run a “supplicant” – software that enables use of 802.1X network authentication protocol, and for wireless, in conjunction with an Extensible Authentication Protocol (EAP). In this context, the switch or AP is called the Authenticator – the device that controls initial network access. In many solutions, the AP forwards these authentication messages on to its Mobility Control (MC) which acts as the authenticator.

It is as if the authenticator says, ”should I grant access to this device? I’ll ask the Authentication Server, which is responsible for making the decision to grant or deny access. So the Switch or AP/MC harvests credentials carried in the 802.1X message, creates a RADIUS message, and sends it to the RADIUS server. RADIUS is both an authentication service running on some server, and it is a protocol used to communicate between Authenticator and Authentication server, used specifically for end user access to network services and apps. To control administrative access to the network components themselves – switches, routers, controllers, a TACACS+ service is often used.

Now the RADIUS server might have a local data store of usernames, passwords, and group memberships, especially in a smaller deployment. In this case, the server checks this data store, validates credentials and sends “access granted” or “access denied” messages to the authenticator, which in turn notifies the endpoint.

However, for larger, more scalable deployments, you have several authentication servers all accessing a common set of backend servers that provide a common data store. This is often a Microsoft server running Active Directory (AD). This sets you up for single sign-on – the same credentials you use to access AD is also used to access the network itself. Nice. Some deployments might use a Lightweight Directory Access Protocol (LDAP) database, and you might also have a Certificate Authority (CA) service for even tighter endpoint and end user authentication.



AAA Services

Authentication Who are you? — Permit or deny access based on credentials/certificates

Authorization What can you do? — Control authenticated user capabilities

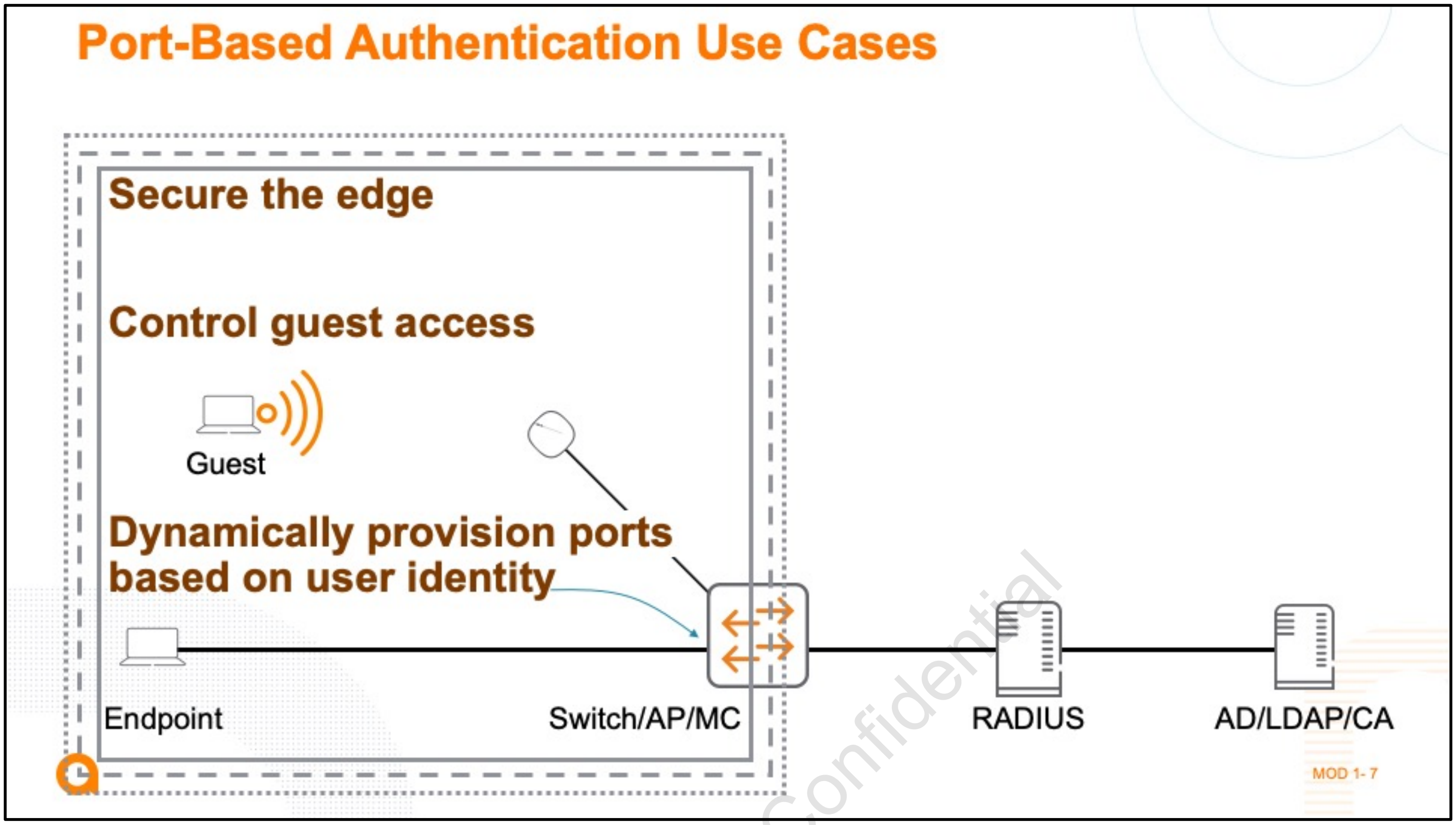
Accounting What did you do? — Auditing, reporting, tracking user activity

So, this is a common Network Access Control (NAC) solution, based around running AAA services on the Authentication Server. So, the three A's of AAA services.

First up is Authentication – who are you? The RADIUS or TACACS+ server permits or denies access based on your credentials. Perhaps your username is Maria, with password Secret123 – if those are valid credentials, access is granted. The user might alternatively submit a digital certificate that includes a valid username and is signed by a trusted certificate authority (CA).

OK, now that you're on the network, what can you do? That's Authorization. User Maria is a member of the marketing group, so she can access all standard business applications, plus the marketing application and database. Other resources are off limits for Maria. You might define a policy to authorize Maria to be on a certain VLAN, with certain Quality of Service (QoS) settings, which the switch port dynamically applies

We have controlled who can access the network, and what they can do. Now we need accounting – what did they do? So we have auditing, reporting, and tracking of user activity – When and from where they logged in. When they logged out, and what resources they accessed during that time. One thing we're trying to eliminate here is "plausible deniability". If the log file indicates that you accessed an HR database at 2:00am, it is difficult to deny that fact.



Port-Based Authentication Use Cases

Secure the edge

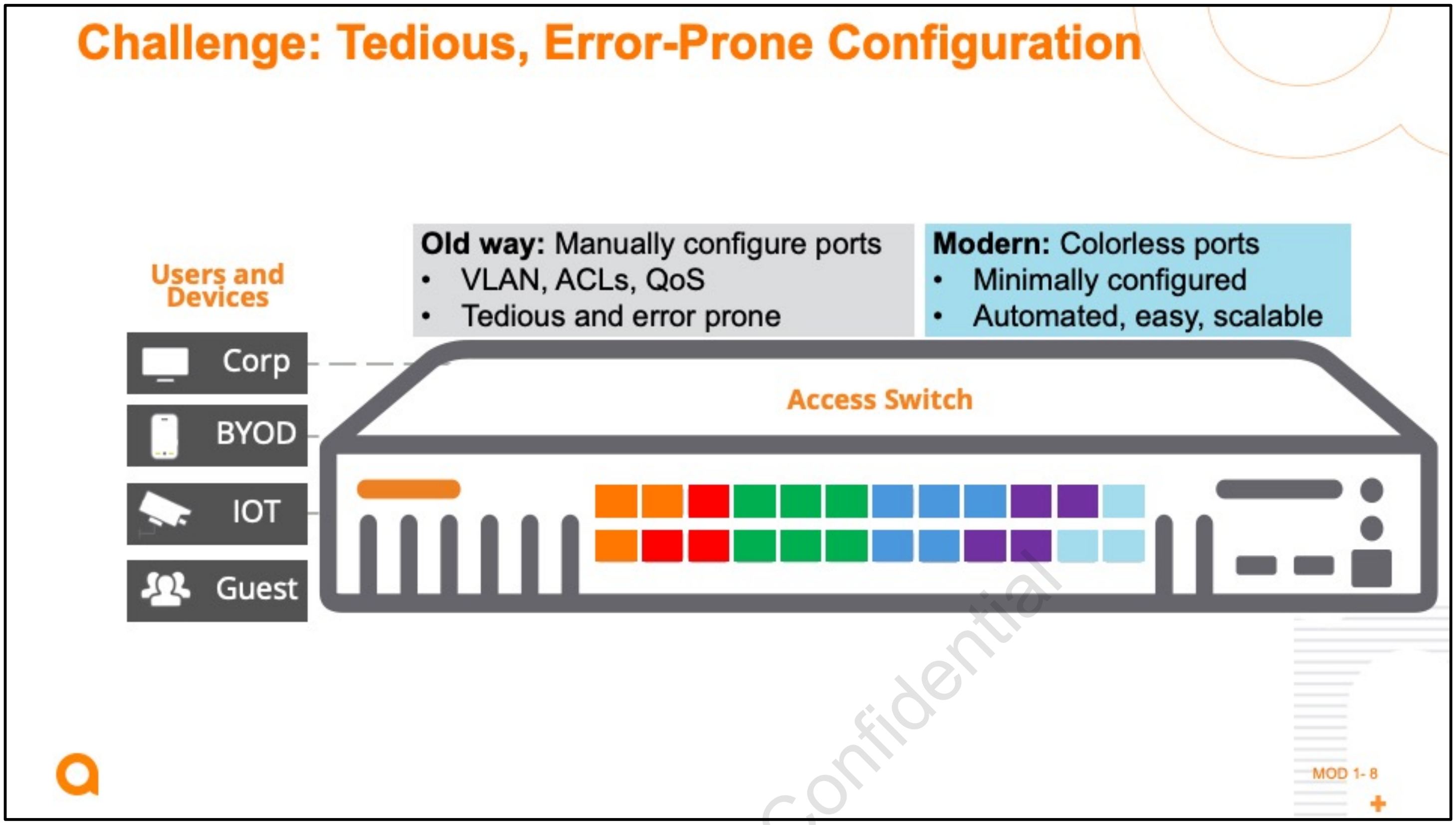
Control guest access

Dynamically provision ports based on user identity

One important use case for port-based authentication it to secure the edge – you must prevent unauthorized users from connecting to open ports. This may be required fo compliance to corporate security policies and government regulations.

Another use case is to control guest access. You must ensure that guests are placed in the proper guest VLAN and must log in for legal and tracking purposes.

Another use case that can really help reduce administrative overhead while improving security is to dynamically provision access ports based on user identity. Let’s take a look.



Challenge: Tedious, Error-Prone Configuration

Old way: Manually configure ports
- VLAN, ACLs, QoS
- Tedious and error prone

Modern: Colorless ports
- Minimally configured
- Automated, easy, scalable

Users and Devices

Corp

BYOD

IOT

Guest

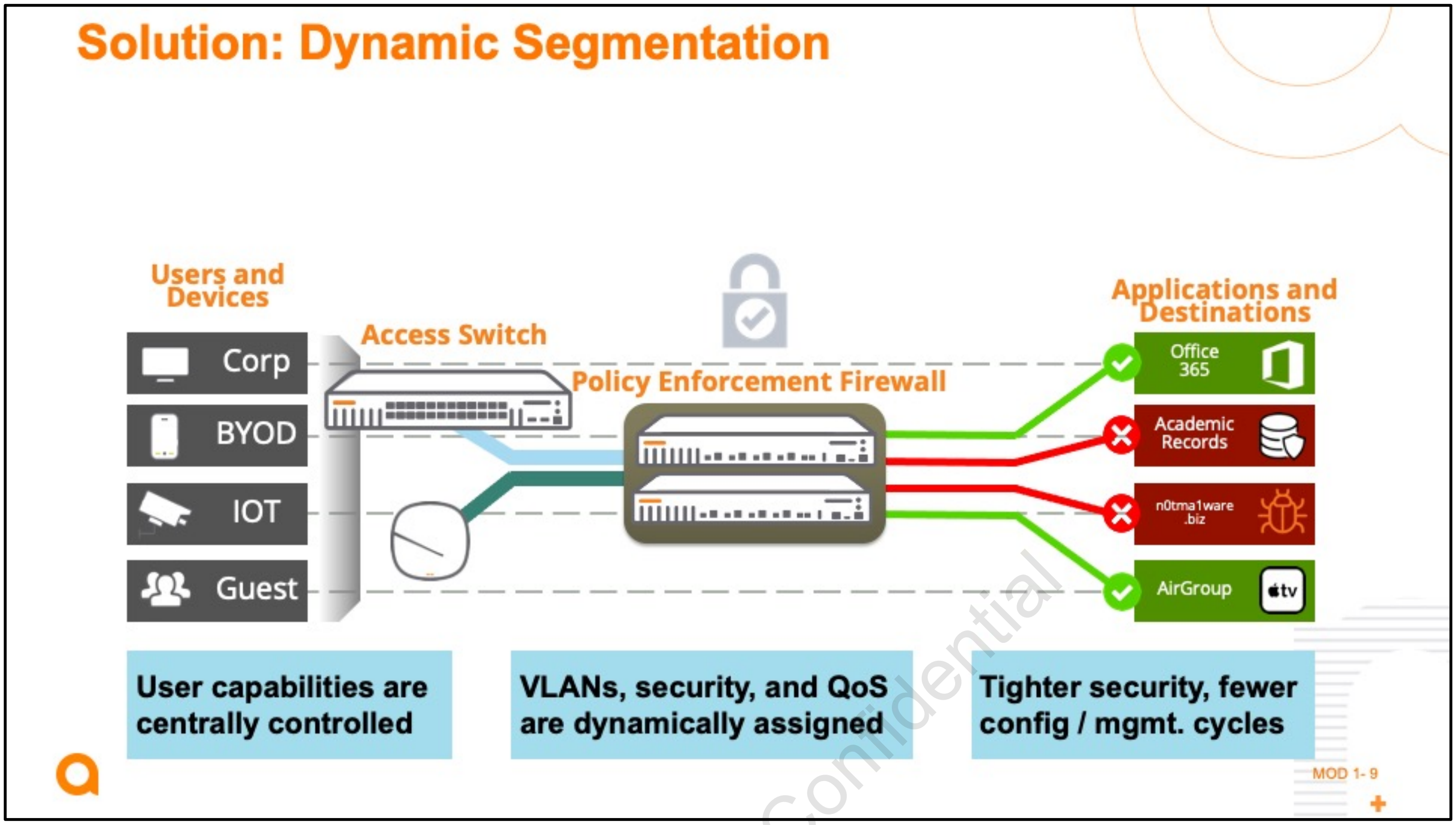
Access Switch

MOD 1- 8

| The old way is to manually configure specific ports for specific VLAN, Access Lists, and Quality of Service – tedious…and you must ensure that everyone connects each device to the correct port – its error prone. Maybe the orange ports are for accounting users, red ports are for IT admin, Green is for sales, Blue is for access points, purple is for printers, and so on.

| Instead, you have a simple collection of minimally configured “colorless ports”.

| Connected users are assigned a role, which Aruba’s ClearPass Policy Manager associates with a particular policy.



Solution: Dynamic Segmentation

Users and Devices

Corp

BYOD

IOT

Guest

Access Switch

Policy Enforcement Firewall

Applications and Destinations

Office 365

Academic Records

n0tma1ware .biz

AirGroup

tv

User capabilities are centrally controlled

VLANs, security, and QoS are dynamically assigned

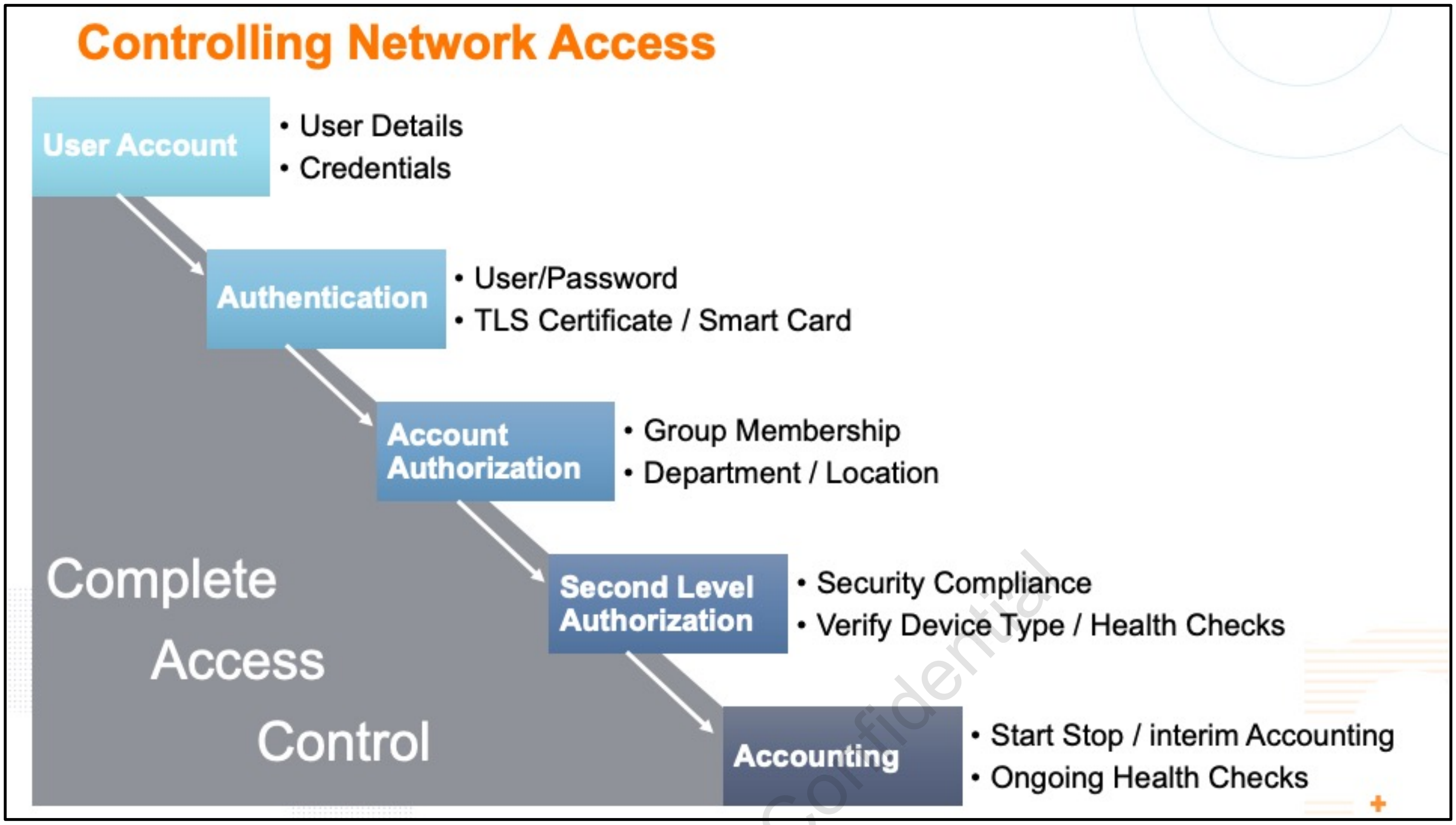
Tighter security, fewer config / mgmt. cycles

MOD 1- 9

That policy is then pushed down to Policy Enforcement Firewall (PEF) features, and a tunnel forms between that PEF device and the user’s connected switch or AP, which permits and denies access to appropriate resources.

|Thus, user capabilities are centrally controlled, and VLAN connectivity, security, and QoS are dynamically assigned. Your network is much more secure, with far less initial configuration and management cycles.

Pretty cool solution, right? But our focus here is on 802.1X and AAA Authentication, to control network access.



Controlling Network Access

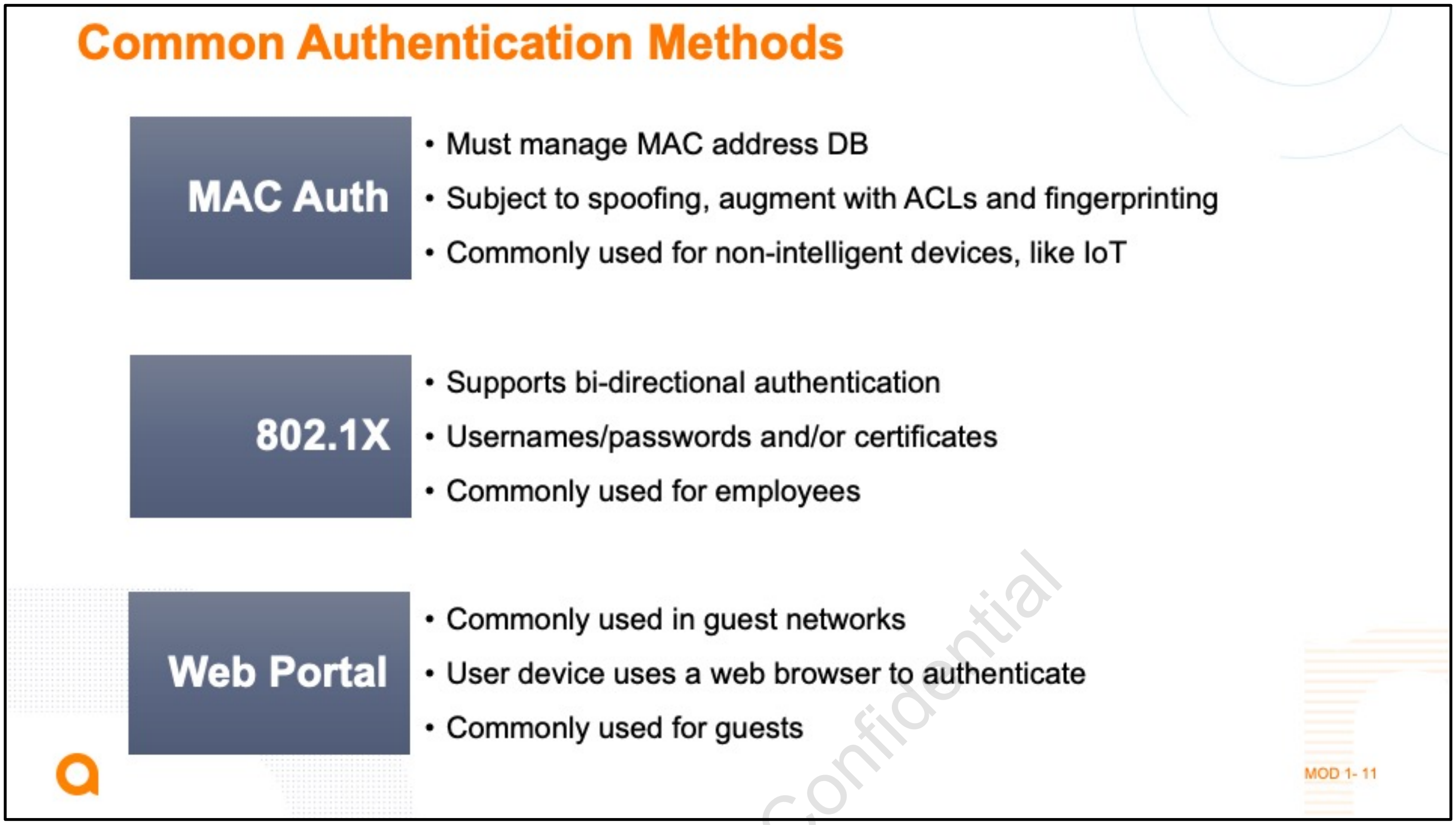
When a user joins the organization, they are given an account. It holds details about their relationship to the organization and their authentication credentials. As you just learned, these credentials are often stored in a Microsoft AD or LDAP database.

| During authentication, the user provides credentials, which the network validates against the data store. This is typically done via a AAA server like ClearPass.

| Then network authorizes the user's account details and client device. The network may grant or deny access based on a complete picture of the client.

| If access is granted, the system continues to monitor client activity using start/stop and interim accounting messages, with ongoing compliance checks.

| By leveraging all available information and implementing sound access rules, ClearPass AAA-based deployment provides a complete access control system, with three common authentication methods.



Common Authentication Methods

MAC Auth

- Must manage MAC address DB
- Subject to spoofing, augment with ACLs and fingerprinting
- Commonly used for non-intelligent devices, like IoT

802.1X

- Supports bi-directional authentication
- Usernames/passwords and/or certificates
- Commonly used for employees

Web Portal

- Commonly used in guest networks
- User device uses a web browser to authenticate
- Commonly used for guests

MOD 1- 11

The authentication process validates user or device identity, to permit or deny network access. Typically, you only wish to permit authorized users with proper network credentials. A MAC address, passphrase, password or certificate can all serve to authenticate users, with credentials pre-configured in some LDAP or AD data store. This is like checking a passport to verify a traveler’s identity.

Click each of the three methods shown in the figure to explore

MAC Auth

With MAC authentication, the Network Access Device (NAS) validates an endpoint MAC address against a list of valid MAC addresses. Endpoint MAC addresses are burned in ROM at the factory, but are copied to RAM during bootup, and this is the MAC address an endpoint adds to Ethernet or 802.11 Wi-Fi frames. This means it is relatively easy for bad actors to change or “spoof” their MAC address to gain unauthorized access.

Since the MAC address in Ethernet and Wi-Fi frames are not encrypted, hackers can use packet capture software to find valid MAC addresses, and then change their MAC address accordingly. Due to these security weaknesses, MAC Authentication is only used for endpoints that cannot support better authentication methods like 802.1X. This includes specialty devices, like some older medical equipment, IoT sensors, and scanners.

To improve overall security with MAC Authentication, use ACLs to strictly limit what devices can access. You can also use device fingerprinting to examine device protocol information, like DHCP and HTTP payload information. Then use this information to identify additional information about the device, like the product, operating system, and other information.

802.1X

802.1X is the most common method used to validate clients on enterprise-class networks, used in conjunction with an Extensible Authentication Protocol (EAP). There are many flavors of EAP, including PEAP, EAP-TLS, EAP-TTLS and others. These EAP types are primarily intended for use with open media such as wireless.

You can think of 802.1X and EAP as “sister protocols” that work as a team. EAP uses 802.1X to transport the user credentials between the user device and the NAS. 802.1X happens at Layer 2, where users are authenticated before an IP address is assigned. This is independent of the Layer 3 protocol. Most NAS solutions will then forward the EAP information onward to a centralized AAA server using the RADIUS protocol.

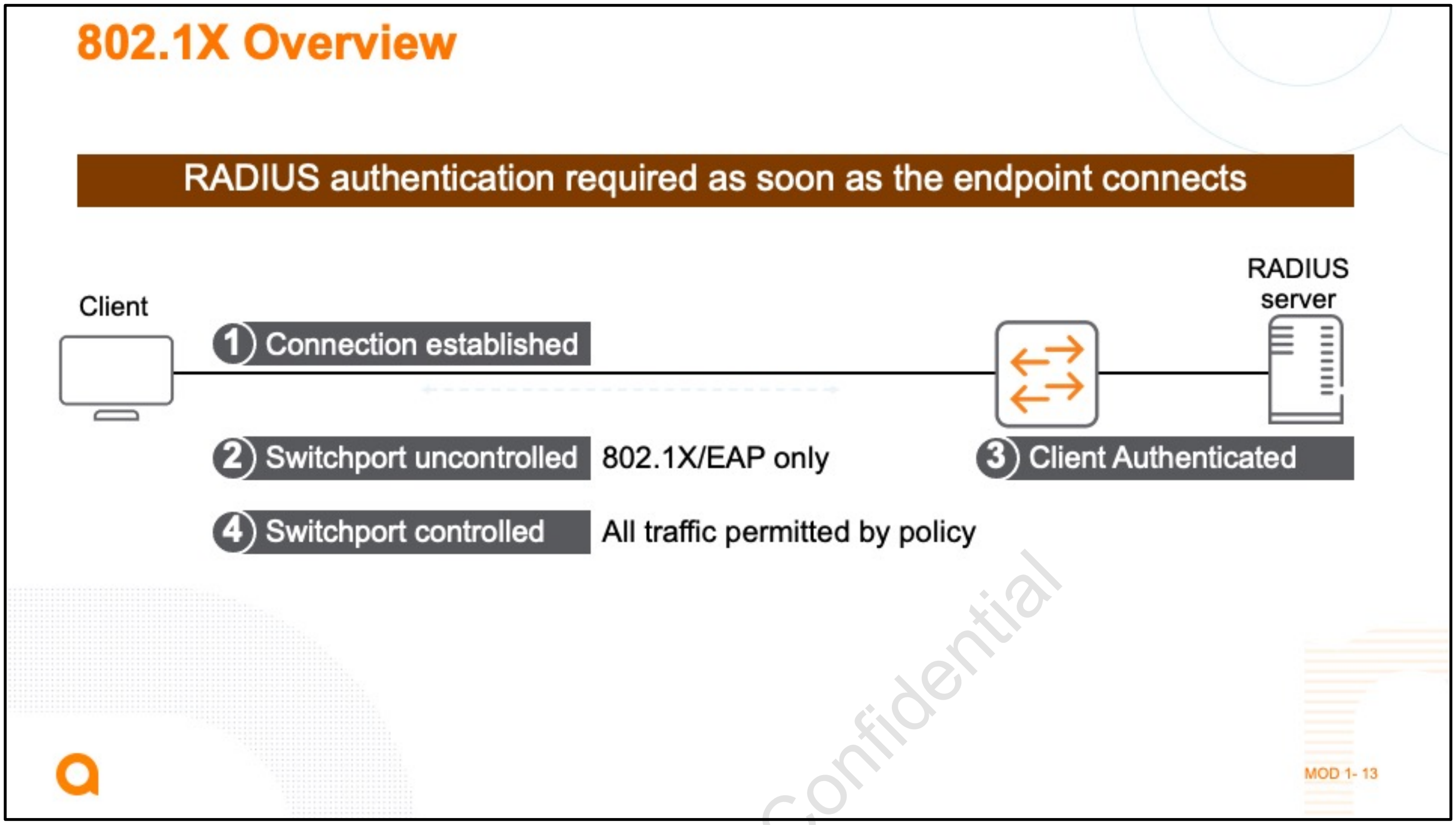
Web Portal

Some systems use Layer 3 authentication such as Captive Portal web authentication or VPN software. Captive portal is commonly used in guest networks to register and restrict their access. Captive portal is web based and thus only requires a web browser on the user device. Given that web is based on the HTTP protocol, only minimal layer 3 access is required - DHCP, DNS, and restricted HTTP to perform the captive portal process.

Basic 802.1X

MOD 1- 12

You will first examine 802.1X, an open standard technology for controlling users' network access based on their identity.



802.1X Overview

RADIUS authentication required as soon as the endpoint connects

802.1X forces users to authenticate as soon as the connection between the switch port and the client comes up. To restrict an unauthenticated user so that client can send only authentication messages, 802.1X divides the port into two logical ports.

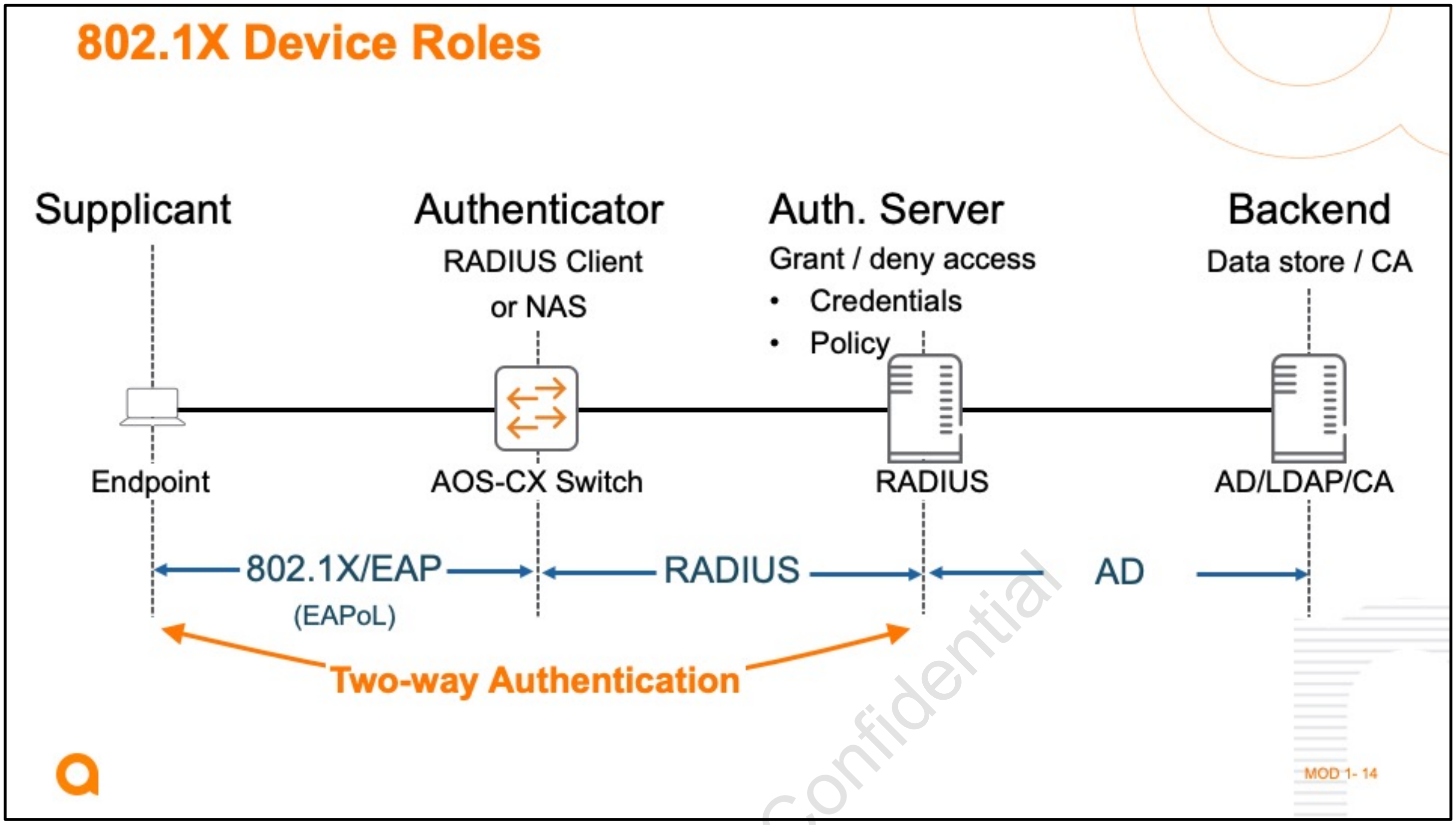
Uncontrolled port

At first, the switchport is an uncontrolled port. This means no process controls whether this logical port is enabled or disabled. It is always active, but only allows 802.1X / EAP packets used for authentication. Successful authentication is required before the controlled port is enabled.

Controlled port

The controlled is typically only enabled upon successful authentication. After authentication, the controlled port can all types of traffic, but you can control this with locally-defined access lists or centrally defined policy. Both sides of the connection control the port based on their assessment of the authentication state. In other words, no unauthorized traffic flows across the link.

This is merely a formal way to ensure that no ingress switch traffic is allowed until you authenticate – except for the 802.1X/EAP traffic required to become authenticated. The uncontrolled (non-authenticated) port does not even allow DHCP or DNS traffic. Authentication happens, and then IP addressing information is assigned.



802.1X Device Roles

Supplicant

Authenticator
RADIUS Client
or NAS

Auth. Server
Grant / deny access
- Credentials
- Policy

Backend
Data store / CA

Endpoint

AOS-CX Switch

RADIUS

AD/LDAP/CA

802.1X/EAP
(EAPoL)

RADIUS

AD

Two-way Authentication

MOD 1- 14

You learned that the supplicant is the client's 802.1X/EAP component, authenticating in response to a challenge from the switch port. But endpoints could receive a challenge before they are fully booted, causing authentication to time out before the user can submit his or her credentials. If this happens, the endpoint can initiate authentication by sending an EAP-Start packet.

You know that the authenticator is the device to which endpoints directly connect. Since these access switches, access points, or firewall devices ultimately control access, they are also referred to as the RADIUS clients or Network Access Servers (NAS). 802.1X EAP messages that traverse a LAN, whether wired or wireless, are also referred to as "EAP over LAN" (EAPoL).

The NAS Authenticator encapsulates these messages in RADIUS packets and send them to the RADIUS Authentication server. Understand that the actual authentication is between the endpoint and the RADIUS server. Think of the NAS as a mediator or translator. The endpoint speaks 802.1X/EAP, the server speaks RADIUS, and the NAS passes packets between the two.

The authentication server can grant or deny access based on several factors:

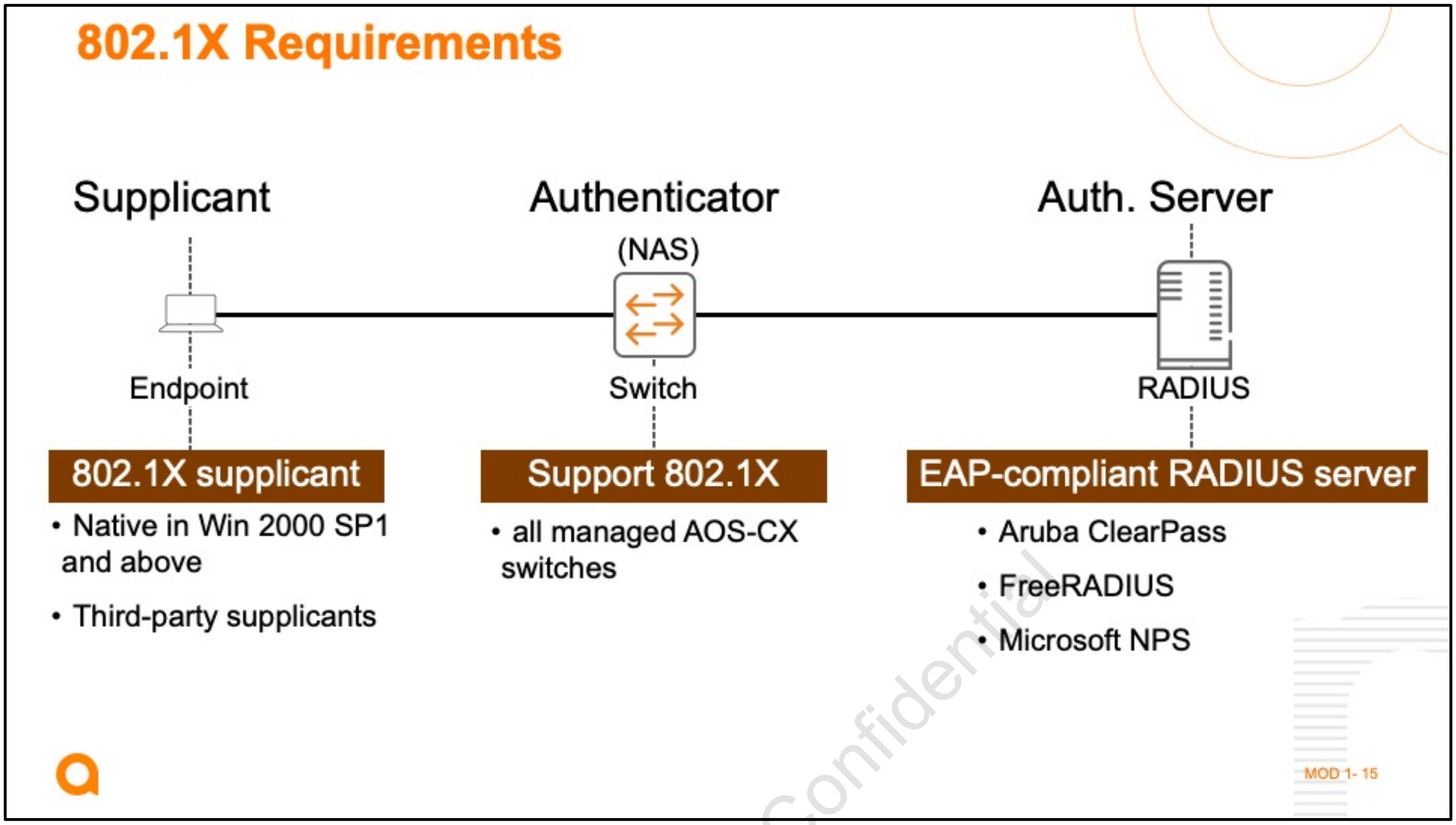
Credentials – username/password, user certificate, and/or machine certificate

Policy and context: Who, what, how, where, when

For example, username John with password Secret123 attempts to login using his corporate laptop via Ethernet connection from Corporate HQ during business hours. He is granted full access. But if he attempts to login using his personal table via Wi-Fi at a café near his home at 1:00am on Saturday, he may only be granted limited access – or no access at all.

Understand that authentication can be in both directions – the network must ensure that the client is valid, and the client must ensure that the network is valid. What if a bad actor sets up a rogue network to which users connect and provide their credentials to the hacker. Bad news!

So when the network says, “Prove your identity – give me your credentials”. The client responds, “ First, you prove your identity!”. The network can provide a server-side certificate, which the client can validate. Then client can provide its username/password, perhaps along with a user and or machine certificate.



802.1X Requirements

Supplicant

Endpoint

802.1X supplicant

- Native in Win 2000 SP1 and above
- Third-party supplicants

Authenticator

(NAS)

Switch

Support 802.1X

- all managed AOS-CX switches

Auth. Server

RADIUS

EAP-compliant RADIUS server

- Aruba ClearPass
- FreeRADIUS
- Microsoft NPS

MOD 1- 15

The figure summarizes requirements for 802.1X on clients, servers, and switches.

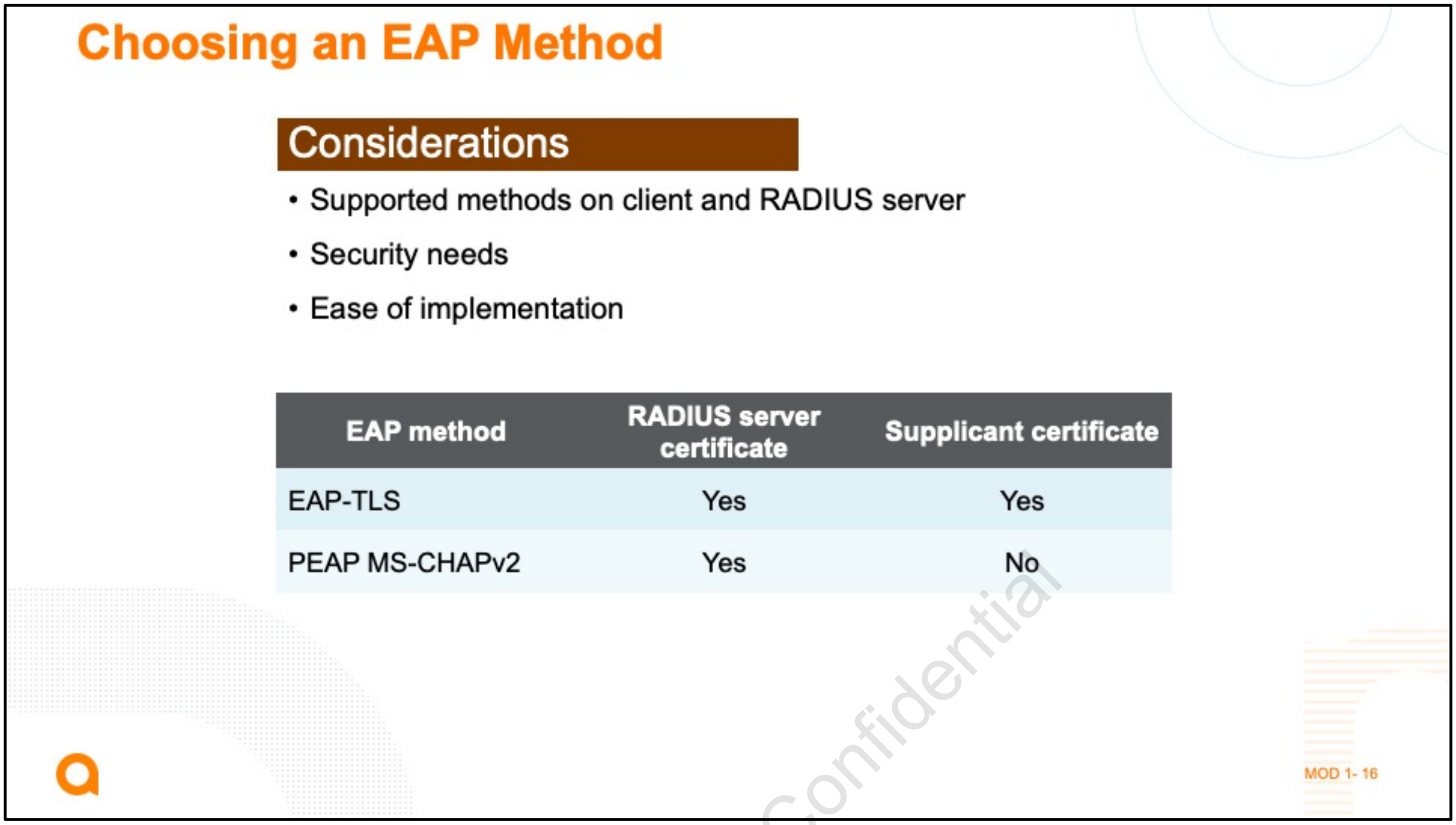
Of course, the endpoint must have an 802.1X supplicant, which comes standard in most modern operating systems like Windows, MacOS, iOS, and others. Some organizations may prefer a third-party supplicant for additional features or personal preference.

RADIUS clients or NASs must support 802.1X of course, as do all managed AOS-CX switches.

And you must have an EAP-compliant RADIUS server.

Aruba ClearPass provides a RADIUS server, as well as other capabilities for monitoring and managing user access. This includes the powerful context features, device profiling, centralized policy management, and more.

You can alternatively use a third-party RADIUS server such as Microsoft Network Policy Server (NPS) or an open-source server such as FreeRADIUS.



Choosing an EAP Method

Considerations

- Supported methods on client and RADIUS server
- Security needs
- Ease of implementation

EAP method	RADIUS server certificate	Supplicant certificate
EAP-TLS	Yes	Yes
PEAP MS-CHAPv2	Yes	No

MOD 1- 16

When choosing an EAP method, you must always consider if both the supplicant and the server support the method. EAP-TTLS tends to be less widely supported. For example, Microsoft Windows does not support TTLS natively; you must install additional software to provide this support. For this and other reasons, EAP-TLS and PEAP MS-CHAPv2 are the most popular choices

NAS devices like the AOS-CX switch do not interpret EAP messages, they simply encapsulate and decapsulate them and forward to the RADIUS server. NAS devices can do this for any valid method.

Note: You can configure local authentication for 802.1X and have the switch act as the authentication server. In this case, you would need to consider the EAP methods that the switch supports. However, this use case is less common and is not covered in this course. Check the Security Guide documentation for more details.

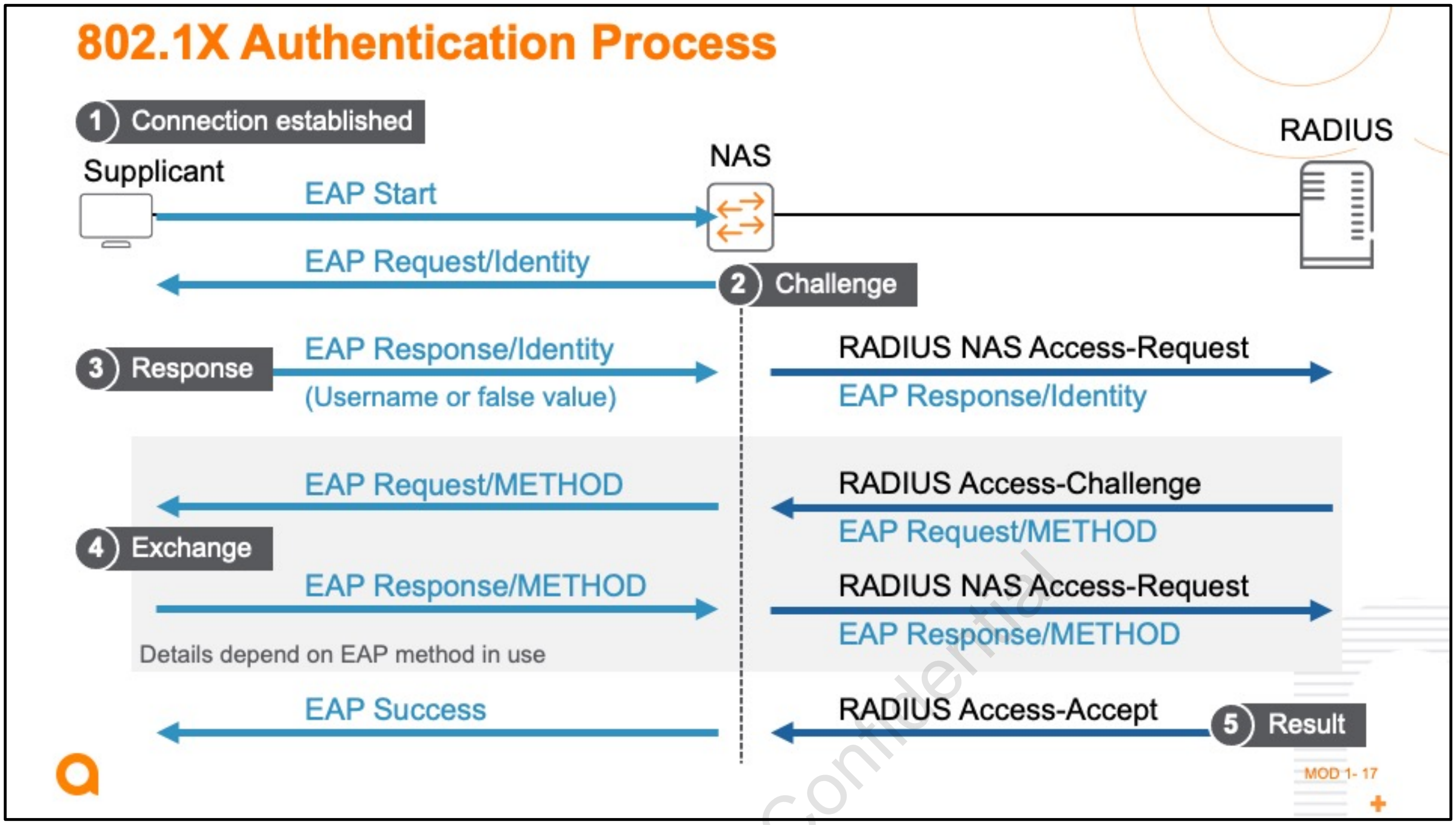
You should also consider security requirements. EAP-TLS tends to be the most secure choice because users authenticate with secure digital certificates. On the other hand, many enterprises prefer PEAP or EAP-TTLS because user names and passwords provide enough security for their purposes. They value of PEAP’s ease of implementation is more important, especially when used an inner method such as MS-CHAPv2, as typical.

EAP-TLS, EAP-TTLS, and PEAP all require that the RADIUS server supply a server-side certificate, so the client can ensure that the network is valid. Many PEAP supplicants do allow you to disable this requirement, but it creates a security risk and is not recommended.

EAP-TLS also requires that each client supply a certificate. Historically, this is the reason why EAP-TLS is both considered more secure, and more difficult to configure, making PEAP a popular alternative. You only need to distribute server-side certificates for a few servers, as opposed to client-side certificates for hundreds or thousands of clients.

However, modern networks now support more automated ways to distribute certificates, and so the difficulty of implementing EAP-TLS is far lower than in the past. The server uses this client-side certificate to authenticate to clients and generate key material for the secure TLS tunnel. You must install a valid certificate, signed by a trusted CA, on the RADIUS server. If you choose to use a self-signed certificate on the RADIUS server, you should install that certificate as trusted on client. Or could configure clients to prompt users to validate an untrusted certificate on the first connection.

Note: EAP-TTLS and PEAP could use TLS as an inner method, which would require a supplicant certificate. However, this is not a typical implementation.



802.1X Authentication Process

1 Connection established

Supplicant

NAS

RADIUS

EAP Start

EAP Request/Identity

2 Challenge

3 Response

EAP Response/Identity

(Username or false value)

RADIUS NAS Access-Request

EAP Response/Identity

EAP Request/METHOD

RADIUS Access-Challenge

EAP Request/METHOD

4 Exchange

EAP Response/METHOD

RADIUS NAS Access-Request

EAP Response/METHOD

Details depend on EAP method in use

EAP Success

RADIUS Access-Accept

5 Result

MOD 1- 17

Let's look at the 802.1X authentication process in detail, to improve your theoretical knowledge and real-world diagnostic abilities.

802.1X carries EAP messages which are used for authentication. EAP defines a flexible framework that supports several authentication methods, to meet your particular environment and policy needs. The figure shows this general framework. As you follow the process, remember that the NAS authenticator relays all messages from the client to the RADIUS server, encapsulating them with RADIUS as necessary. The vertical line underneath the switch shows the point at which the authenticator encapsulates the messages.

| The client and switch port establish an Ethernet connection.

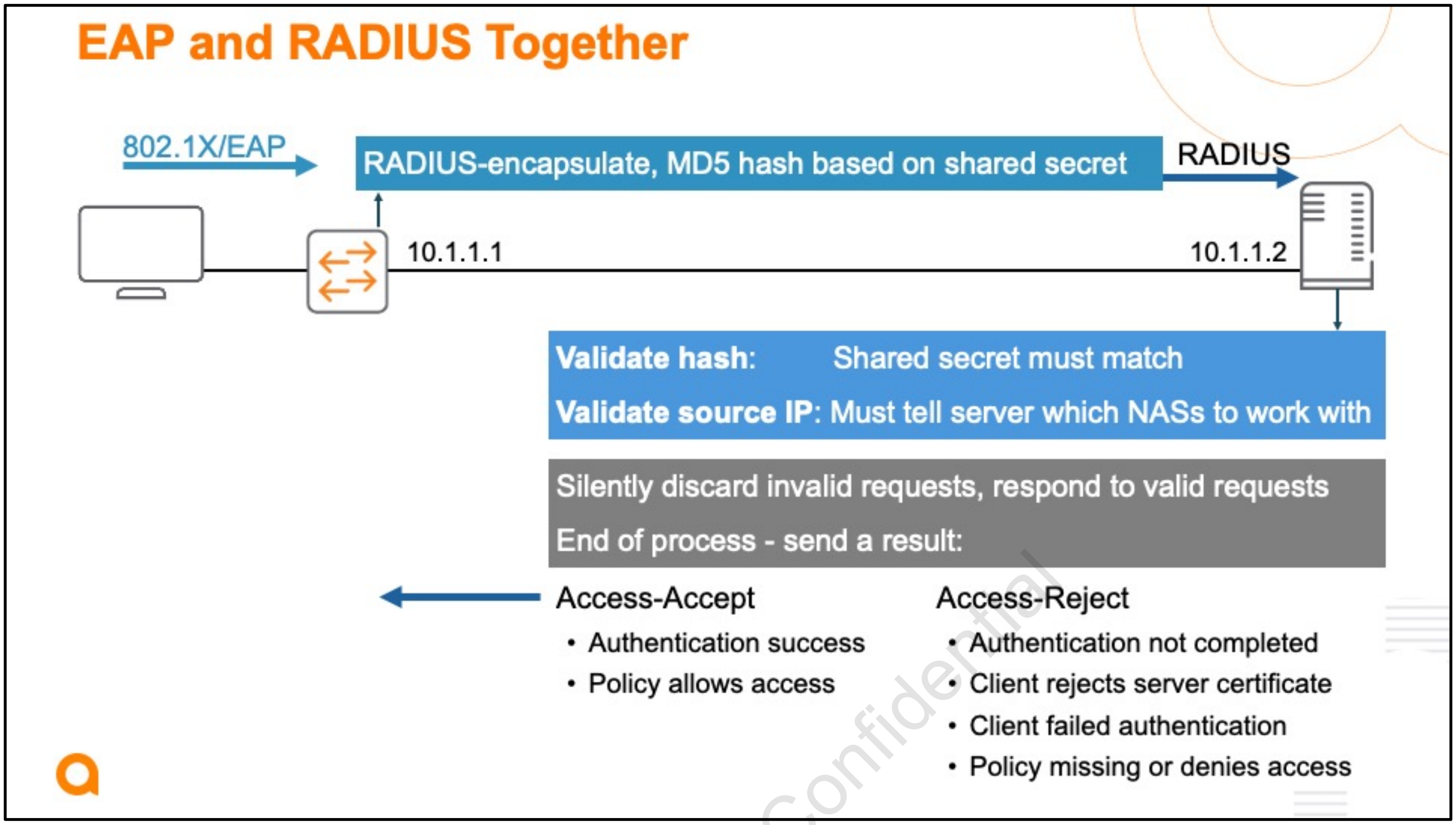
| The switchport is controlled, blocking all traffic except EAP, and sends an EAP Request/Identity challenge. This initiates authentication and asks the client only to identify itself – do not send any other information.

|The client responds with an EAP Response/Identity packet, which typically includes a username. This is encapsulated in a RADIUS message, and forwarded to the RADIUS authentication server.

The NAS authenticator includes the user's identity in all future frames so that the server can keep track of which EAP messages belong to which user. The user's identity also marks any accounting frames for the connection - important for companies with rigorous auditing requirements. Depending on the EAP method, though, the username submitted at this point might not be the actual username. Some methods have the supplicant submit a false name at this point and the true name later in the process after a secure, encrypted session is established. This protects the username from hackers.

| The server initiates an EAP method for exchanging credentials. Message exchange details, flow, and authentication credentials depend on the EAP method.

| Based on information received in the exchange, the authentication server grants or denies access and informs the switch. The switch then either activates the connection and transmits an EAP-Success or leaves the connection deactivated and transmits an EAP-Failure.



EAP and RADIUS Together

802.1X/EAP

RADIUS-encapsulate, MD5 hash based on shared secret

RADIUS

10.1.1.1

10.1.1.2

Validate hash: Shared secret must match

Validate source IP: Must tell server which NASs to work with

Silently discard invalid requests, respond to valid requests

End of process - send a result:

Access-Accept

- Authentication success
- Policy allows access

Access-Reject

- Authentication not completed
- Client rejects server certificate
- Client failed authentication
- Policy missing or denies access

Now let’s look at how the authenticator (switch) relays messages between the client supplicant and RADIUS server. You must configure the NAS and RADIUS server to work together. You must configure the NAS with the RADIUS server’s IP address, so it knows where to send packets. And you must configure the RADIUS server with the IP addresses of acceptable NAS devices, so it knows who to accept requests from. Each must be configured with a matching shared secret – a simple alphanumeric string of your choice.

The authenticator encapsulates a supplicant’s EAP message in a RADIUS Access-Request packet. It also hashes the packet with MD5, using the RADIUS shared secret as the key. It adds the hash to a message authenticator field. The authenticator then transmits the packet to the RADIUS server. This communication is between the NAS IP and the Server IP address - the client does not need an IP address or any network access.

The RADIUS server checks that the message has arrived from a valid network device (sometimes called a Network Access Server [NAS] or RADIUS client); that is, the switch IP address must match a network device address configured on the server. The RADIUS server also checks the message authenticator field by hashing the packet with its shared secret. If the request is not from a known network device with a correct shared secret, the RADIUS server does not respond to the message – it silently discards the request.

If the switch is known and the secret is correct, the server decapsulates the EAP messages and formulates the appropriate responses, which it encapsulates in RADIUS Challenge messages. These messages are also authenticated with the message authenticator field.

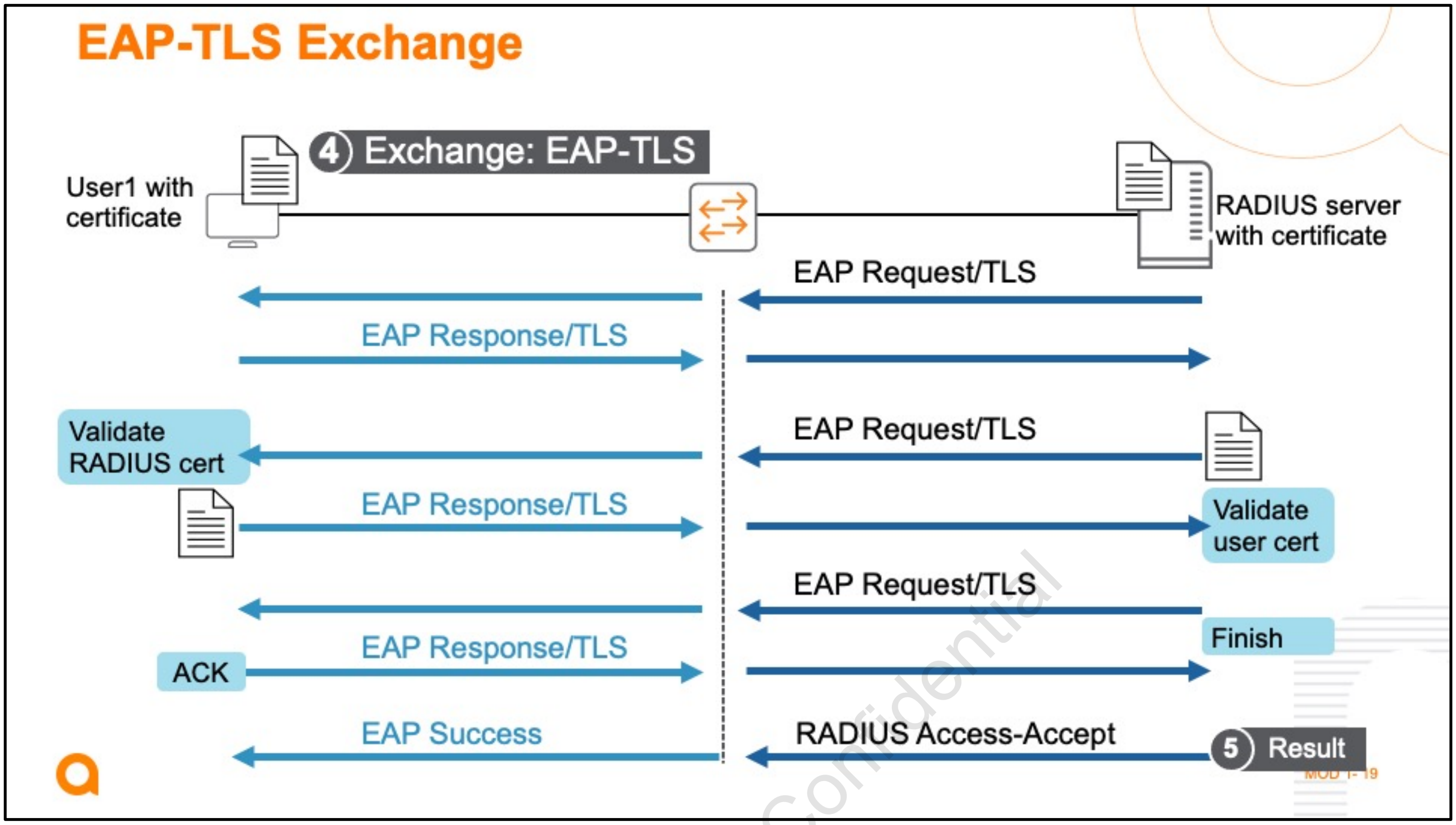
At the end of the process, the RADIUS server sends either:

An Access-Accept message—Two requirements are met:

- The client has successfully authenticated.
- The policy that applies to this access request allows access.

An Access-Reject message—One of these situations occurred:

- The client could not complete authentication, whether because the RADIUS server does not trust the authenticator, because the client and server do not support compatible EAP methods, or some other problem.
- The client does not trust the RADIUS server's certificate, so the server failed authentication.
- The client failed authentication.
- The policy that applies to this access request does not allow access.
- No policy applies to this access request.



EAP-TLS Exchange

4 Exchange: EAP-TLS

User1 with certificate

RADIUS server with certificate

EAP Request/TLS

EAP Response/TLS

Validate RADIUS cert

EAP Request/TLS

EAP Response/TLS

Validate user cert

EAP Request/TLS

EAP Response/TLS

ACK

Finish

EAP Success

RADIUS Access-Accept

5 Result

MOD 1- 19

As mentioned earlier, step 4 in the EAP/RADIUS process is where users prove their identity.

The server's EAP Request/METHOD packet starts the process and indicates the method that the server requires. You can configure most RADIUS servers to select methods according to conditions such as user identity and location.

If the station supports the requested method, it continues the exchange. Otherwise, the station sends an EAP NAK packet, which can suggest a different method. The server, if it supports the alternative method, may then initiate the exchange with the new method. The station and the RADIUS server must agree upon the method for the process to continue.

The server sends their certificate to the user, which validates it against a local store of acceptable certificates. The EAP Request/TLS and EAP Response/TLS packets include information such as:

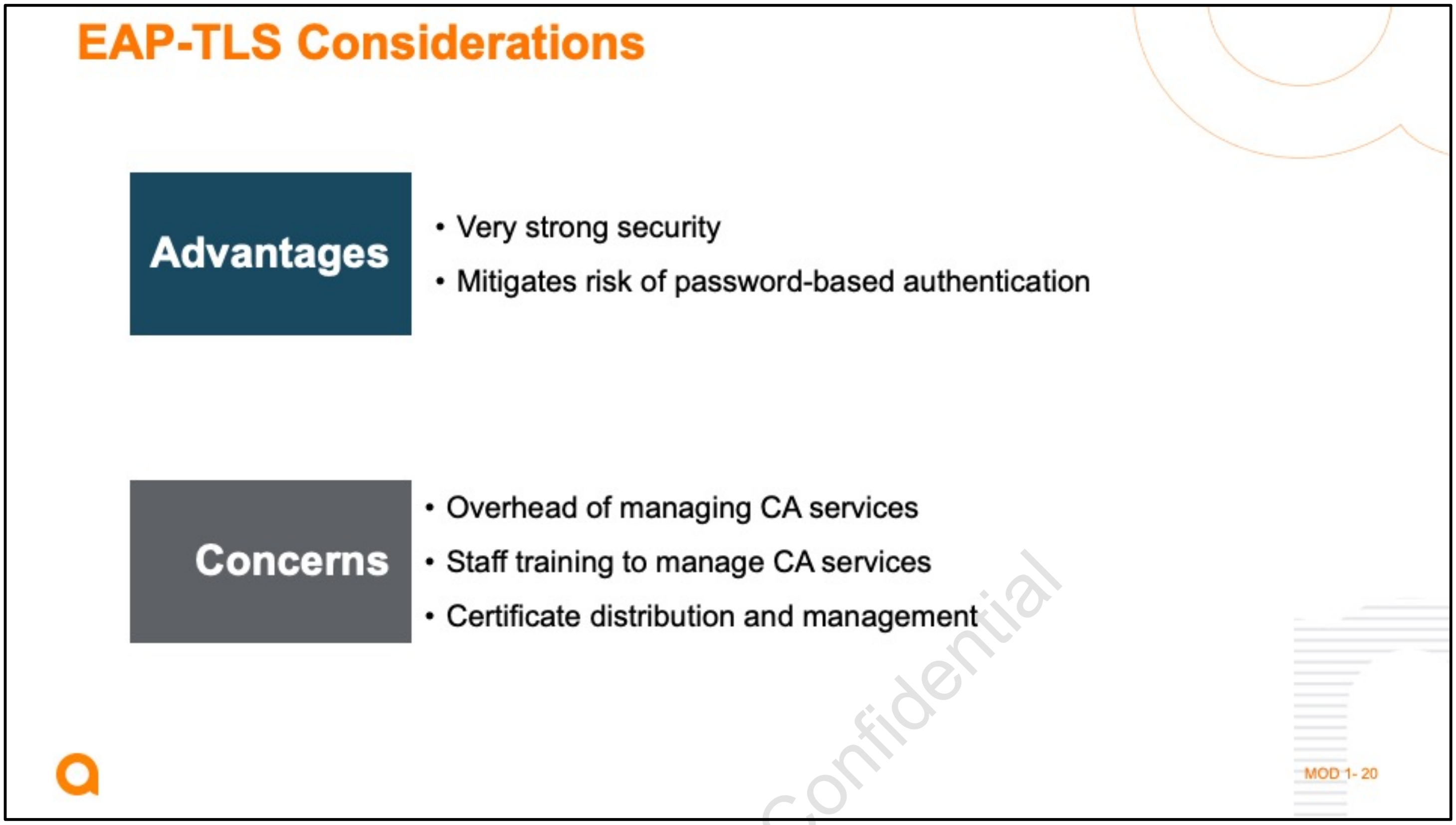
Digital certificates and certificate verifications

Supported encryption suites

Values for generating encryption keys (not the keys themselves)

The digital certificate includes the device's identity, and it is signed by a Certification Authority (CA). The RADIUS server validates the certificate submitted by the user, and the supplicant validates the RADIUS server’s certificate. Validation occurs through asymmetric encryption with public and private keys, which is beyond the scope of this course. The validation process also includes checking the CA signature, so the certificate must be signed by a CA that the device is configured to trust.

After the server validates the user certificate, it sends a message to indicate a successful finish, and the client responds. Recall that after the EAP exchange successfully completes, step 5 occurs: the RADIUS server sends a RADIUS Access-Accept.



EAP-TLS Considerations

Advantages

- Very strong security
- Mitigates risk of password-based authentication

Concerns

- Overhead of managing CA services
- Staff training to manage CA services
- Certificate distribution and management

MOD 1- 20

The figure summarizes the advantages and concerns related to EAP-TLS. Please click on each box in the figure to explore.

Advantages

EAP-TLS is one of the most secure EAP methods because it provides mutual authentication with Public Key Infrastructure (PKI) digital certificates. Digital certificates rely on extremely strong asymmetric keys and are validated by trusted Certificate Authorities (CA). The digital certificate is typically stored on a laptop or a smart card, providing stronger security. Although you can set up requirements to force users to create “stronger,” and therefore more secure, passwords, you cannot prevent users from telling people their password or writing it on a paper displayed in plain sight. Generally, it is harder for users to leak certificates inappropriately; systems administrators might, for example, implement security measures to prevent users from copying certificates.

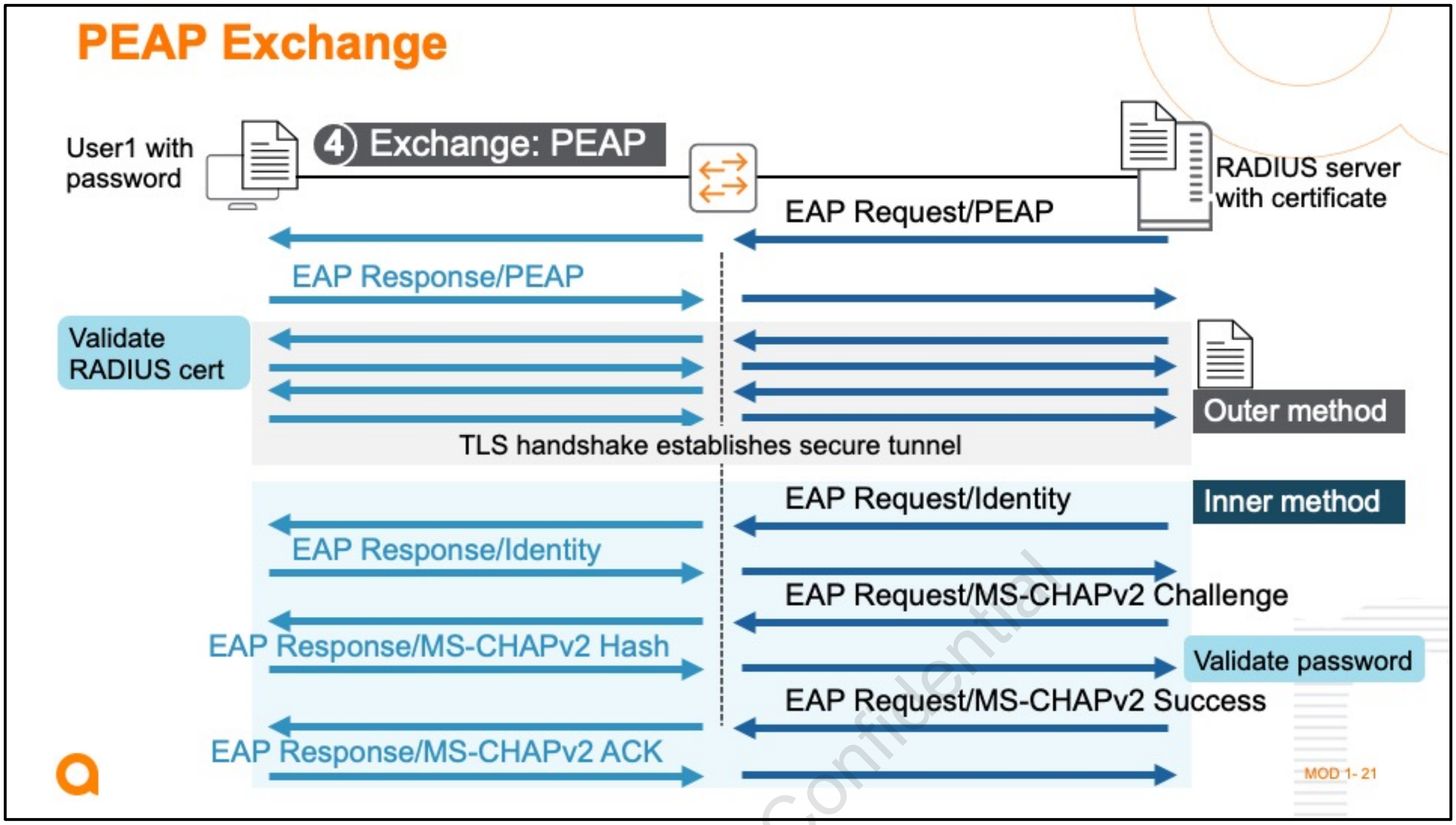
Hackers often run dictionary attacks on passwords, phish for passwords, or use other social engineering techniques to get users to leak their passwords. Although no credentials can be totally secure—malware might be able to steal certificates, for example—digital certificates are more resistant to many common attacks.

Someone who steals a laptop or smart card might be able to gain access to the certificates within, but they are usually protected by an additional layer of security such as a password, which the user must enter before the supplicant can use the certificate. In addition, the user can immediately report the theft, allowing you to revoke the compromised certificate.

Concerns

EAP-TLS is not vulnerable to eavesdropping and replay attacks that affect less secure EAP methods such as EAP-MD5, but security comes at the cost of purchasing and managing the digital certificates—substantially more expensive than managing passwords. Maintaining a large number of certificates requires specialized software and trained IT staff.

In addition, in environments that have multiple operating systems, installing certificates on all clients can be difficult. Bring Your Own Device (BYOD) environments can make managing certificates even more difficult. However, Aruba ClearPass BYOD onboarding capabilities can help companies surmount these challenges. ClearPass Onboard can provide new user devices with limited access and helps users



PEAP Exchange

User1 with password

4 Exchange: PEAP

RADIUS server with certificate

EAP Request/PEAP

EAP Response/PEAP

Validate RADIUS cert

Outer method

TLS handshake establishes secure tunnel

EAP Request/Identity

Inner method

EAP Response/Identity

EAP Request/MS-CHAPv2 Challenge

EAP Response/MS-CHAPv2 Hash

Validate password

EAP Request/MS-CHAPv2 Success

EAP Response/MS-CHAPv2 ACK

MOD 1- 21

The figure starts at step 4 of the 802.1X authentication process, this time showing a PEAP MS-CHAPv2 exchange.

Step 1—outer method

Like EAP-TLS, TTLS and PEAP use a TLS handshake to generate encryption keys and negotiate a tunnel secured by those keys. Only the server authenticates itself with a digital certificate during this exchange, but you must ensure that the clients trust the CA that signed the server certificate. The encryption keys for the tunnel are derived from the public key in this certificate.

Step 2—inner method

The client authenticates itself in the second step; it uses a weaker—and so more easily implemented—authentication method, protected by the secure tunnel established during the outer method exchange.

The RADIUS server sends a new EAP Request/Identity packet, and the supplicant responds with their identity. The server then begins the inner authentication method.

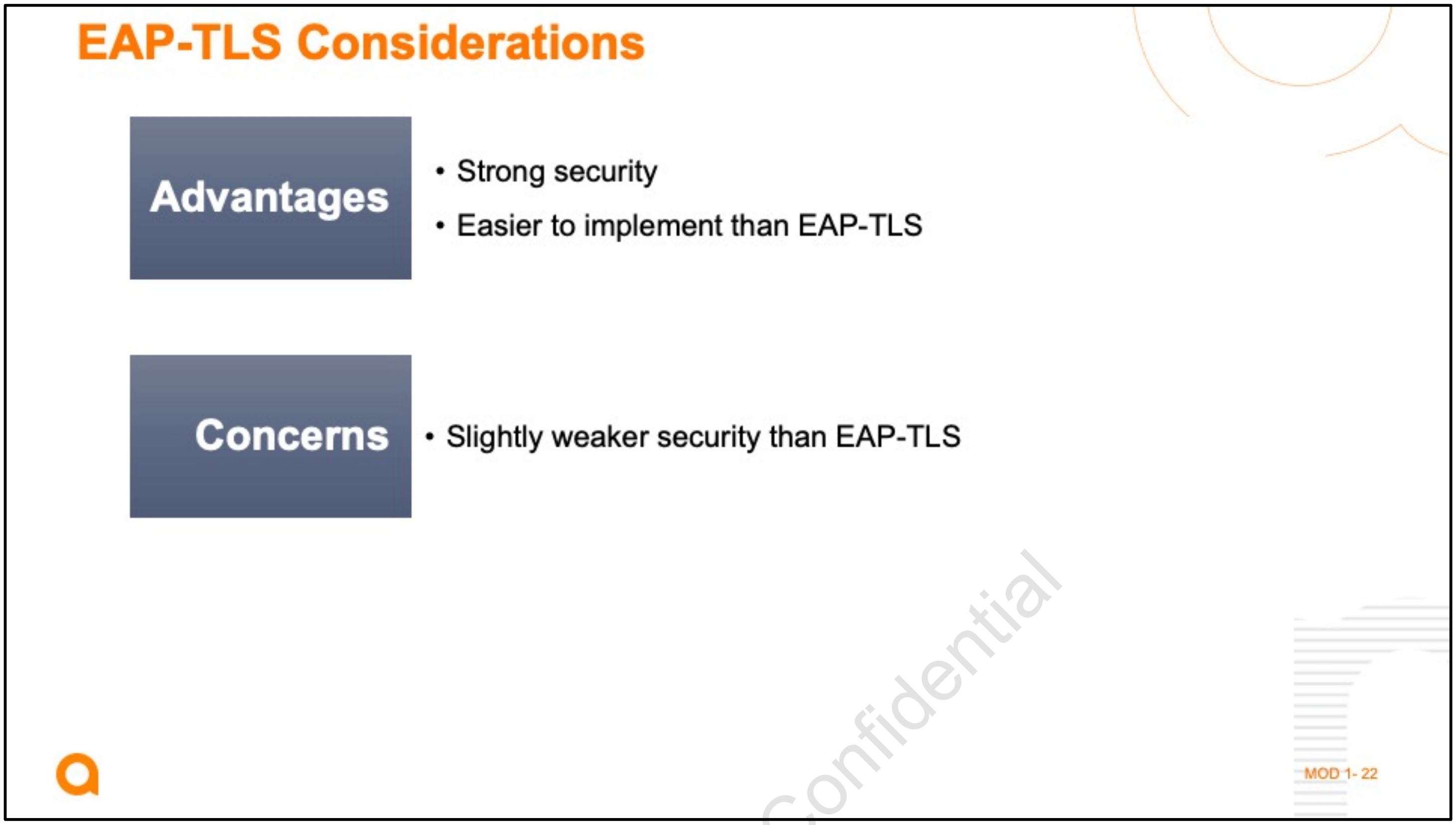
For the inner authentication method, TTLS can use a weaker EAP method, such as EAP-GTC, or a legacy RADIUS method, such as CHAP, PAP, or Microsoft CHAP variants (MS-CHAP v1 or MS-CHAP v2). PEAP supports methods such as MS-CHAP v2, EAP-GTC, and TLS

Because Windows native supplicants support PEAPv0 with MS-CHAP v2, this is by far the most prevalent EAP method - explained here.

The RADIUS server sends a challenge. The supplicant responds with a hash of the password. The server validates that the hash is correct, and, if so, it sends a message indicating the success. The supplicant responds.

The figure shows only the PEAP part of the complete 802.1X authentication process. After these messages, the tunnel is taken down, and the RADIUS server sends a RADIUS Access-Accept to the switch.

Note: For 802.1X on wireless connections, whether implemented through EAP-TLS, EAP-TTLS, or PEAP, the authentication server sends the authenticator the encryption material. The wireless client and authenticator continue to use encryption to secure the connection. On Ethernet connections, however, the client and switch port to do not encrypt traffic in this way.



EAP-TLS Considerations

Advantages

- Strong security
- Easier to implement than EAP-TLS

Concerns

- Slightly weaker security than EAP-TLS

MOD 1- 22

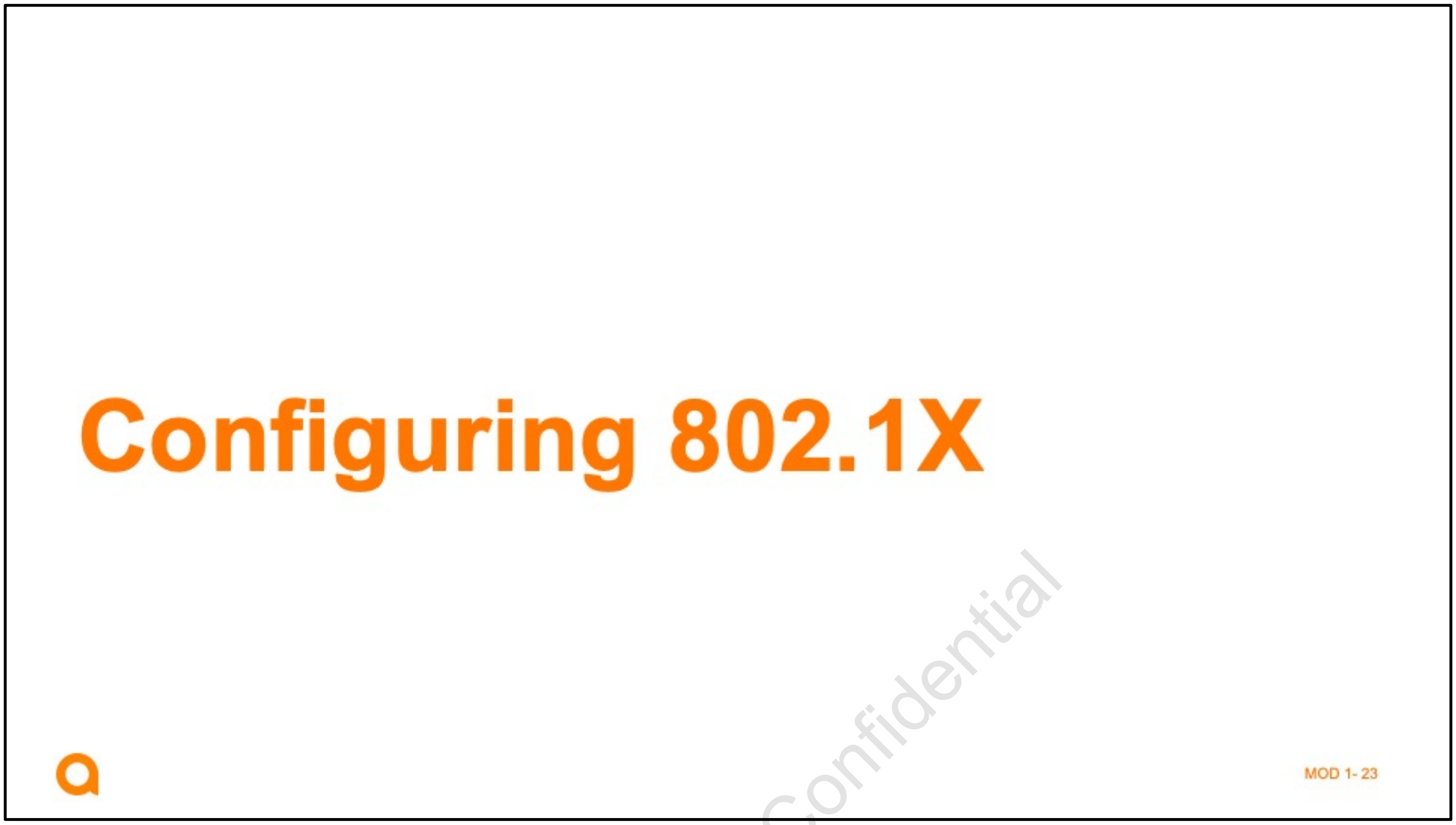
Advantages

Like EAP-TLS, EAP-TTLS and PEAP provide strong, mutual authentication. Because TTLS and PEAP use encrypted tunnels to secure usernames and passwords (rather than require digital certificates on stations), you can implement these methods more easily than you can TLS.

TTLS has one unique benefit: it always protects the username. Depending on how PEAP is implemented, the username might be transmitted in plain text in the initial EAP/Response Identity packet, allowing a hacker to detect the user's identity and possibly lock the user out of his or her account.

Concerns

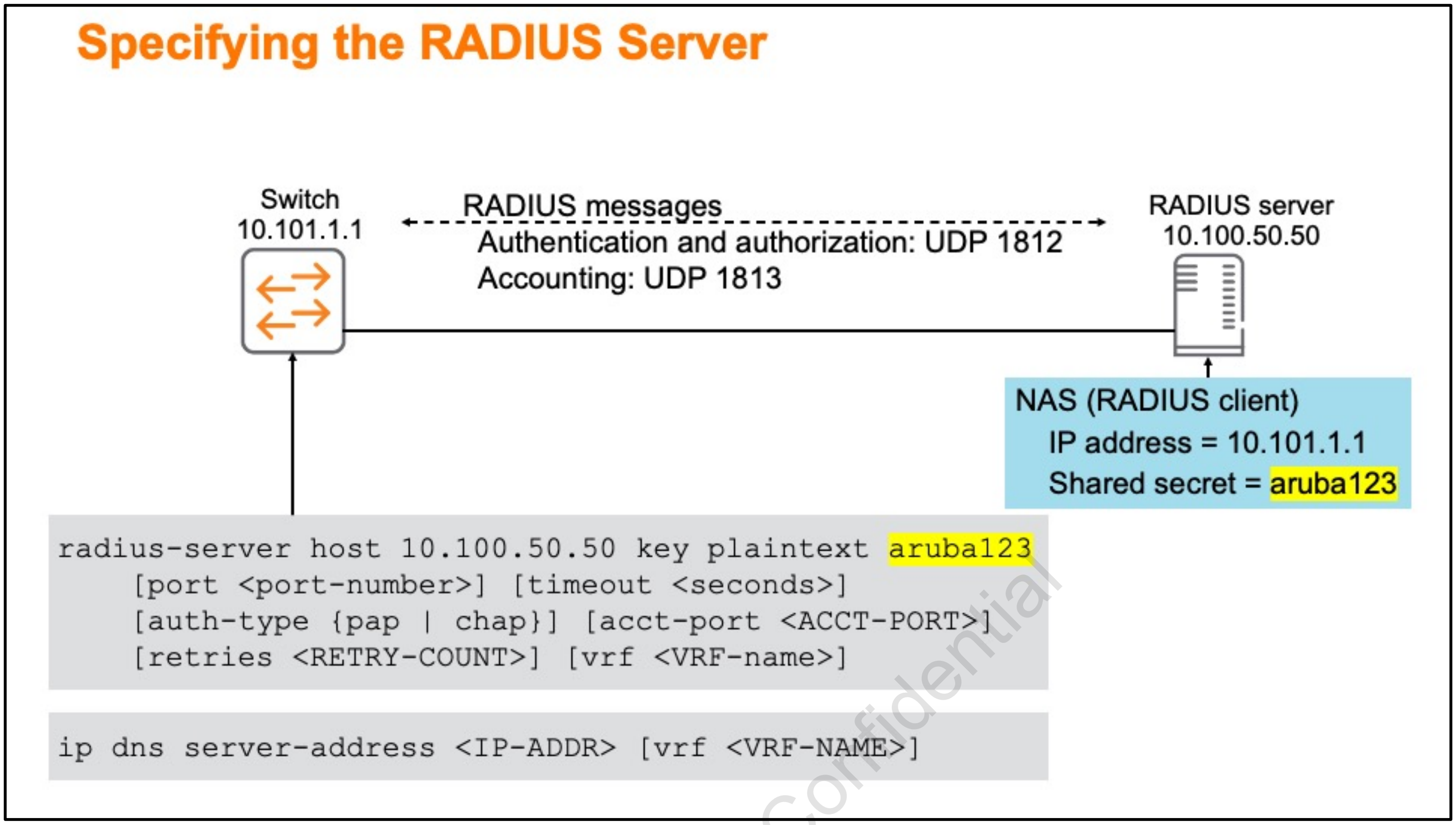
As compared to EAP-TLS, EAP-TTLS and PEAP have the disadvantage of a weaker authentication method.



Configuring 802.1X

MOD 1- 23

And now you learn to configure 802.1X.



Specifying the RADIUS Server

```
radius-server host 10.100.50.50 key plaintext aruba123
    [port <port-number>] [timeout <seconds>]
    [auth-type {pap | chap}] [acct-port <ACCT-PORT>]
    [retries <RETRY-COUNT>] [vrf <VRF-name>]
```

```
ip dns server-address <IP-ADDR> [vrf <VRF-NAME>]
```

Before you enable 802.1X on a switch, you must tell the switch who to send RADIUS messages to and optionally, how to send them. Use the syntax shown in the figure, which refers to the RADIUS server by IP address. You can also use the Fully Qualified Domain Name (FQDN) if preferred.

If you define a FQDN name for the RADIUS server, you'll have to define a DNS server to resolve the name to an IP address

The switch sends a RADIUS Access-Request to this IP address whenever it needs to complete any type of RADIUS authentication. By default, the switch uses the RADIUS standard UDP port 1812 for the authentication messages. After the RADIUS server has authorized a user for access, the switch can continue to communicate information about the user session in accounting packets sent to port 1813.

Note: You can also include the option auth-port and acct-port options in the radius-server host command to change the authentication and accounting ports. Other options are documented in the switch Security guide.

Global key or shared secret

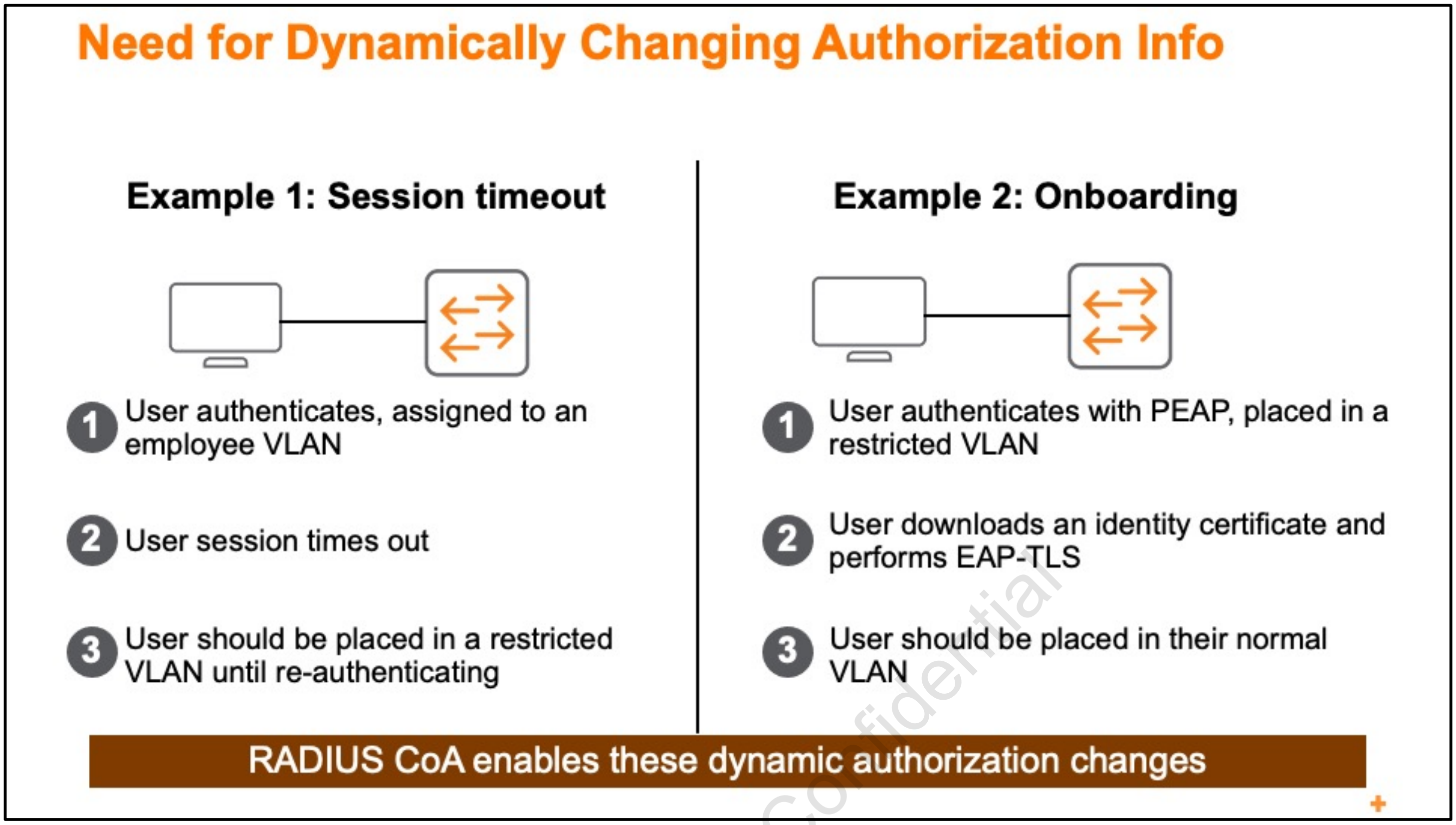
The key option in the radius-server host command sets an individual shared secret for communicating with the RADIUS server. The RADIUS server has the switch IP address configured on it as a RADIUS client (sometimes called Network Access Server (NAS) or, on ClearPass, a network device). The shared secret configured for the switch on the RADIUS server must match the key configured for the RADIUS server on the switch. Otherwise, the RADIUS server will not respond to the switch's messages.

Setting the key for the individual server is optional if you set a global key. An individual key configured for a host overrides the global key, but if you do not set an individual key, the RADIUS server uses the global key. You might choose to use the global key if you have multiple RADIUS servers that use the same shared secret. The command for setting a global key is:

```
Switch(config)# radius-server key plaintext <secret-key>
```

Defining a RADIUS host with a key will override the global key setting.

Note: RADSEC (RADIUS with IPSec) is supported by AOS-CX switch. To configure this, see the Security guide.



Need for Dynamically Changing Authorization Info

Example 1: Session timeout

1. User authenticates, assigned to an employee VLAN
2. User session times out
3. User should be placed in a restricted VLAN until re-authenticating

Example 2: Onboarding

1. User authenticates with PEAP, placed in a restricted VLAN
2. User downloads an identity certificate and performs EAP-TLS
3. User should be placed in their normal VLAN

RADIUS CoA enables these dynamic authorization changes

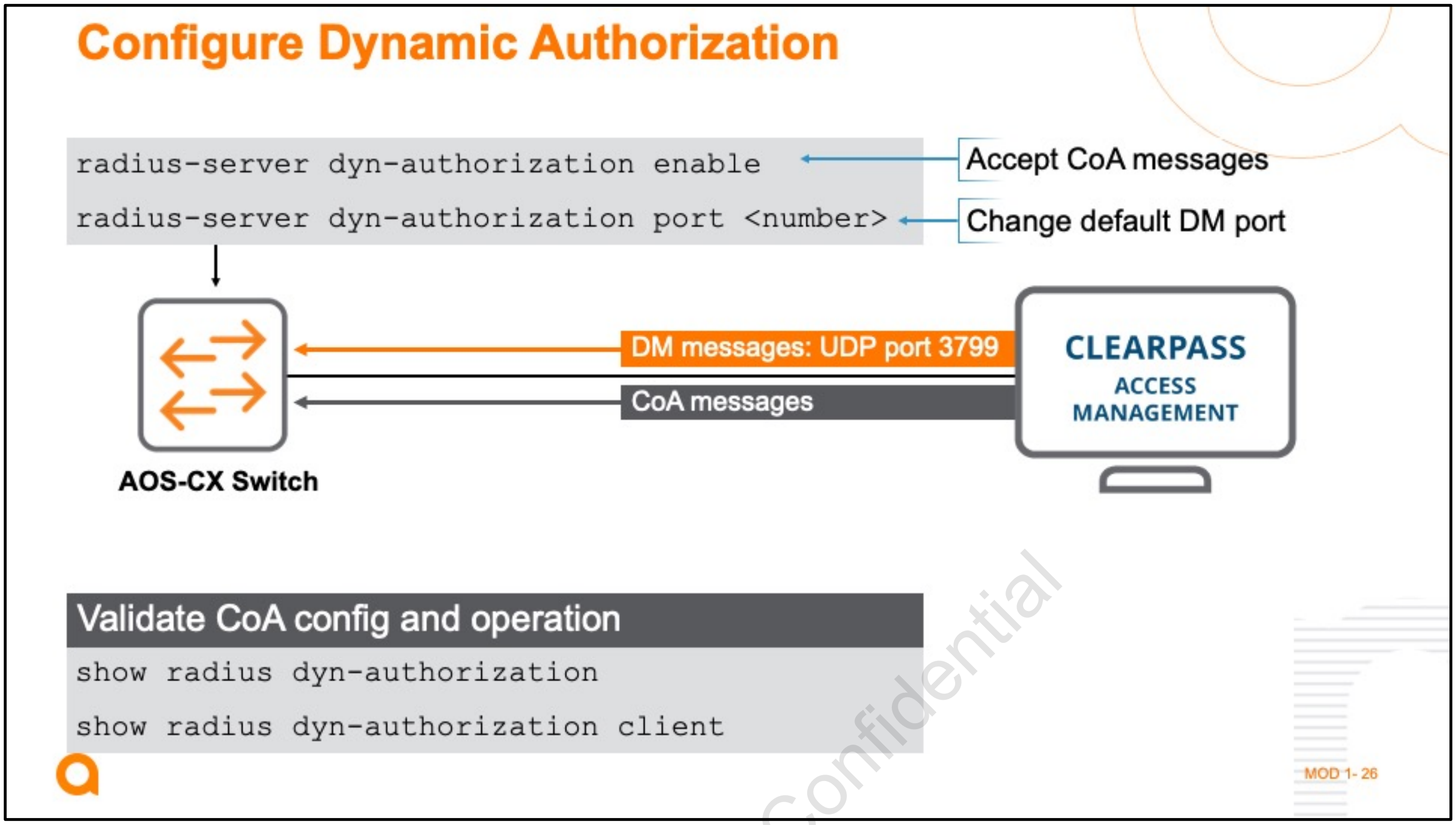
The examples illustrate situations where a user has authenticated, but their authorization information must change based on what is happening in the network.

In the first example, the user authenticates and is assigned their normal VLAN. However, once the user's session times out, they should be placed in a restricted VLAN until they re-authenticated.

| In the second example, a company originally deployed PEAP for 802.1X, but is migrating to certificates. If a user connects and uses PEAP, they should be placed in a restricted VLAN that allows them to perform onboarding. Once the user is onboarded and has downloaded a root certificate and a new identity certificate, they should re-authenticate using EAP-TLS and be placed in their normal VLAN.

There are many examples where a company needs to change authorization information based on what devices are doing in the network: using the authorization information originally assigned to the user when they authenticated can create problems based on these network changes. RADIUS supports a Change of Authorization (CoA) that allows a RADIUS server to push new authorization information to a network device and have it use it for a user's session.

Important: Most networking devices ignore RADIUS CoA messages from an AAA server unless you explicitly enable it. This is true of both ArubaOS and AOS-CX switches.



Configure Dynamic Authorization

```
radius-server dyn-authorization enable
radius-server dyn-authorization port <number>
```

Validate CoA config and operation

```
show radius dyn-authorization
show radius dyn-authorization client
```

When you use Aruba ClearPass as the RADIUS server, it is recommended that you add a few settings to ensure correct operation between the AOS-CX the ClearPass RADIUS server.

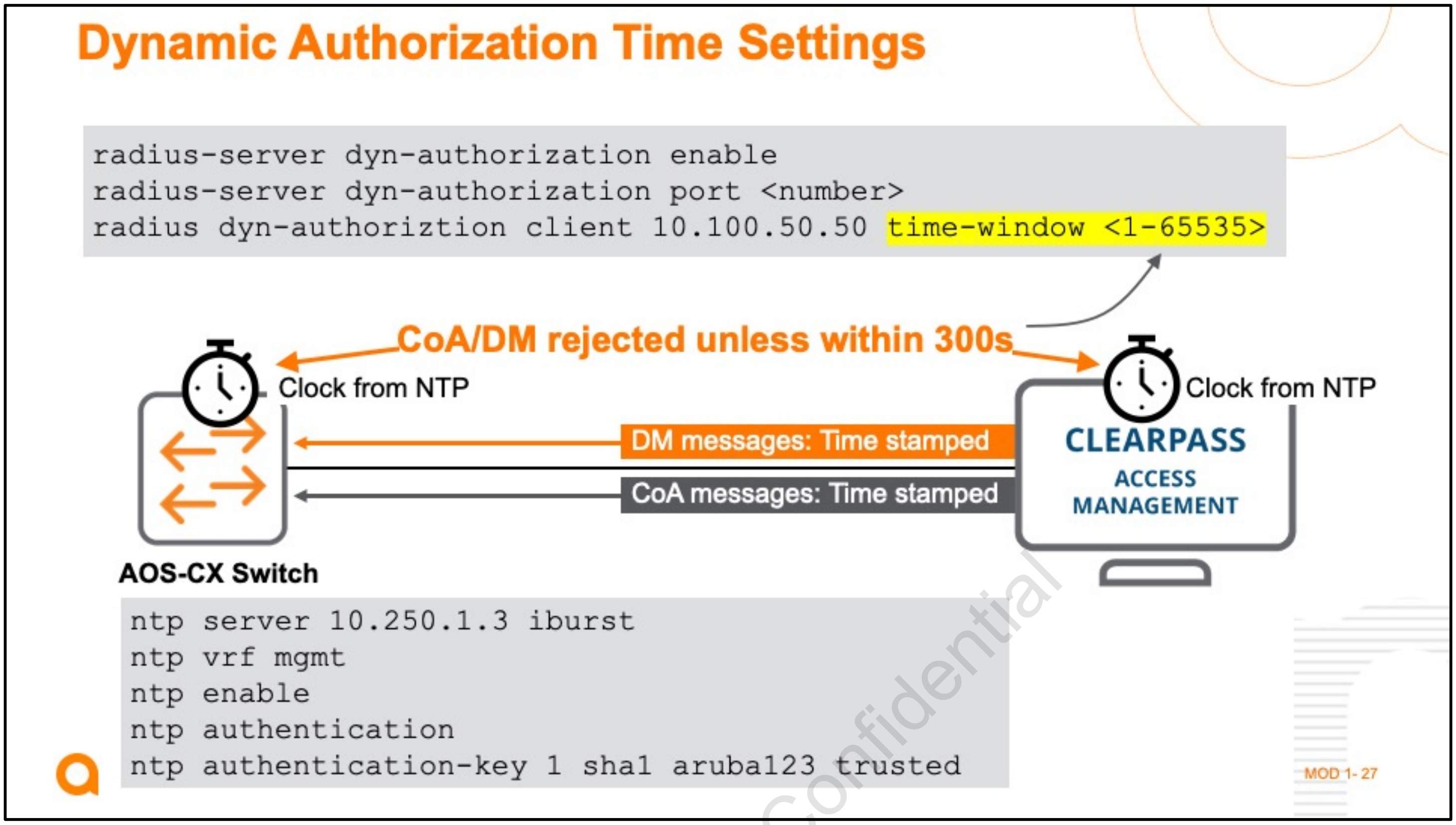
Enable acceptance of Change of Authorization (CoA) messages

Aruba ClearPass can assess endpoints on an ongoing basis and change an endpoint's authentication status or settings. To make these changes, ClearPass sends disconnect messages (DMs) and CoA messages. For the switch to accept these messages, you must enable dynamic authorization for the RADIUS globally on the switch, as shown in the figure. Without this option, some solution components might work, but others will fail.

If dyn-authorization is enabled, the switch accepts these messages on the default UDP port 3799, which is also the default port on ClearPass. You can change that setting, if necessary, in your environment, as shown in the figure.

If you are using a different RADIUS server from ClearPass, you should check whether the server needs to send DMs and CoA messages.

Use the show radius dyn-authorization command to verify the configuration and operation of CoA for all servers, and restrict the view to one server with the show radius dyn-authorization client command.



Dynamic Authorization Time Settings

```
radius-server dyn-authorization enable
radius-server dyn-authorization port <number>
radius dyn-authoriztion client 10.100.50.50 time-window <1-65535>
```

CoA/DM rejected unless within 300s

```
ntp server 10.250.1.3 iburst
ntp vrf mgmt
ntp enable
ntp authentication
ntp authentication-key 1 sha1 aruba123 trusted
```

MOD 1- 27

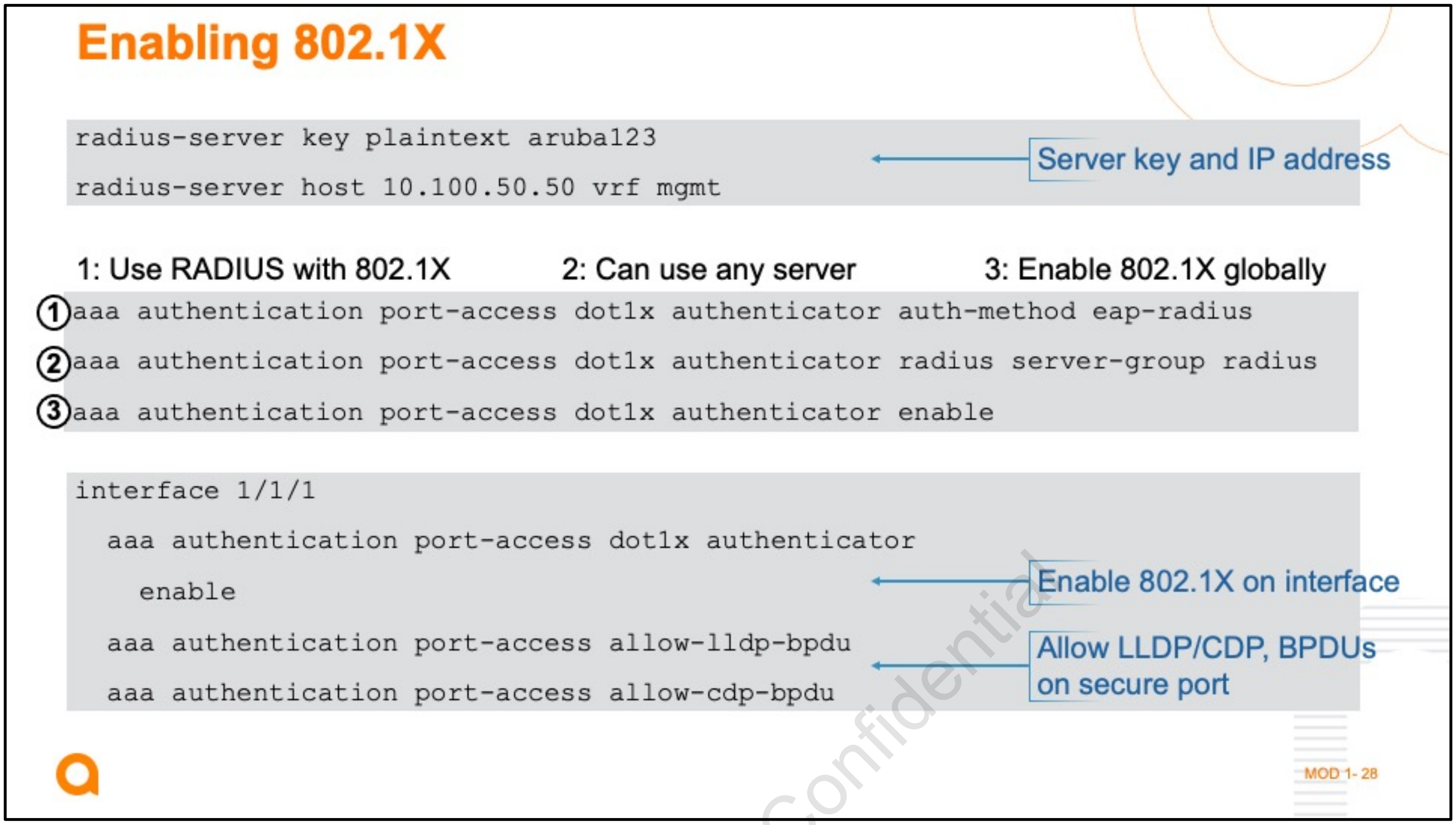
When an AOS-CX switch is configured to allow CoA and DM messages, it checks these messages to determine if they are current. The message should include a time stamp for the event that caused the RADIUS server to change the user status. The switch only accepts the message if the time between the time stamp and the current time on the switch falls within its RADIUS time window.

The default time window is 300 seconds. You can adjust the time window as shown in the figure, optionally specifying a vrf.

To ensure that the switch can accept the CoA and DM messages, make sure that the switch has the correct time. It is best practice to configure NTP on the switch and the RADIUS server and have them take their time from the same clock.

You must also determine whether the RADIUS server always includes a timestamp with its DM and CoA messages. ClearPass does not.

Unlike the older ArubaOS switches, you cannot disable the time window check. Instead set up NTP, as shown in the figure. For the explanation of these commands, see the Fundamentals Guide.



Enabling 802.1X

```
radius-server key plaintext aruba123
radius-server host 10.100.50.50 vrf mgmt
```

Server key and IP address

1: Use RADIUS with 802.1X 2: Can use any server 3: Enable 802.1X globally

```
①aaa authentication port-access dot1x authenticator auth-method eap-radius
②aaa authentication port-access dot1x authenticator radius server-group radius
③aaa authentication port-access dot1x authenticator enable
```

```
interface 1/1/1
  aaa authentication port-access dot1x authenticator
    enable
  aaa authentication port-access allow-lldp-bpdu
  aaa authentication port-access allow-cdp-bpdu
```

Enable 802.1X on interface

Allow LLDP/CDP, BPDUs on secure port

MOD 1- 28

To configure 802.1X, first specify the RADIUS server IP address and key.

Next, configure RADIUS with 802.1X, specify that it can use any defined server, and then enable 802.1X globally.

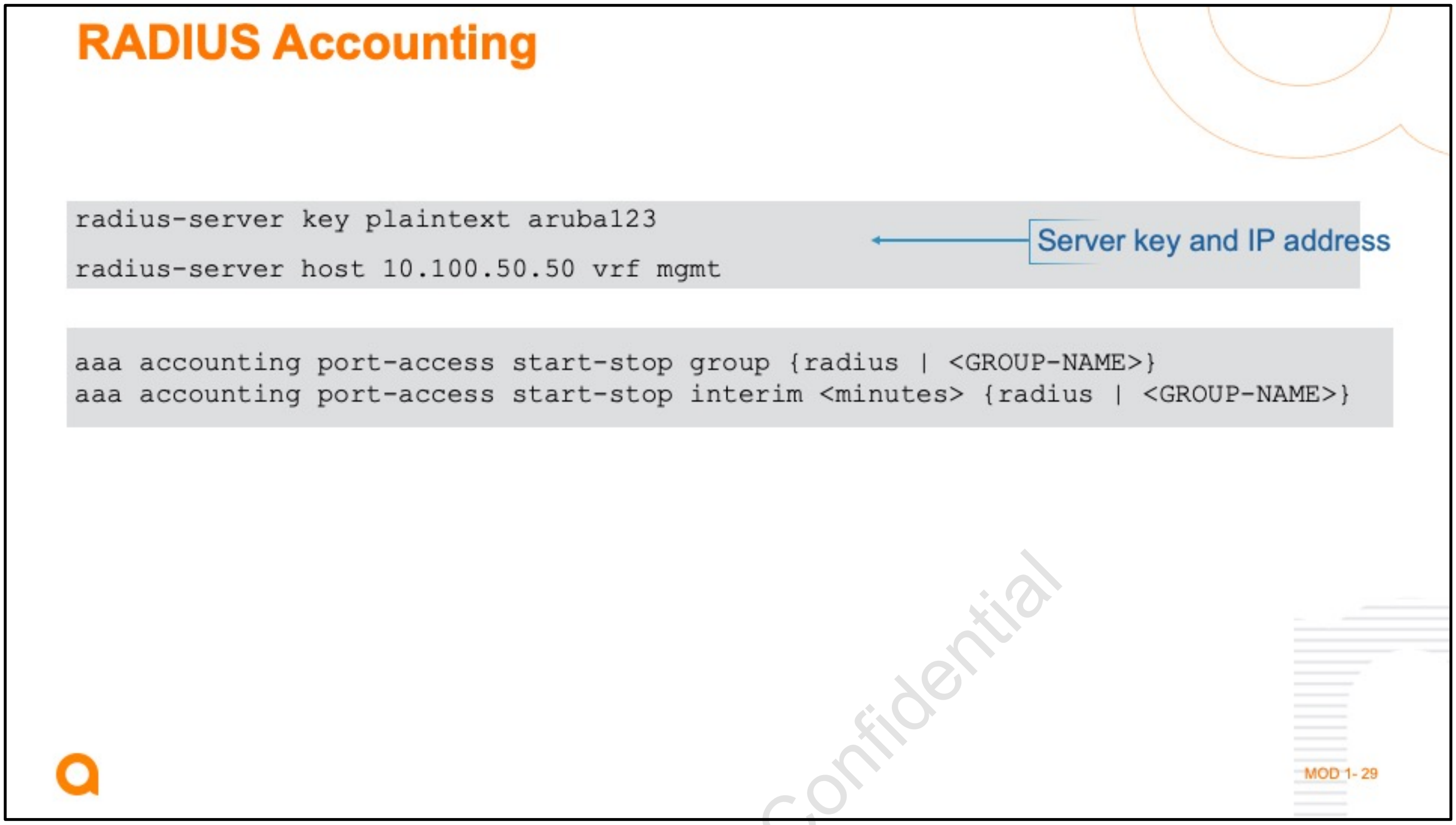
The first server defined with the radius-server host command will be used, by default. Soon you'll see how to restrict the use of particular RADIUS servers by creating a group.

Last, you enable 802.1X on a per-interface basis, as shown in the figure.

Now the switch will act as described earlier. It will keep the port closed to all non-EAP traffic and send EAP messages to trigger an 802.1X supplicant on the connected device to authenticate. If the user succeeds in authenticating to the RADIUS server, the switch opens the port for normal traffic. If the user fails, the port remains closed.

If you need to allow LLDP/CDP or BPDU packets on a secure port, add the optional configuration shown in the figure.

This configuration might be necessary for devices like a phone or camera where LLDP (or Cisco's CDP) must be exchanged for the device to operate correctly prior to authentication occurring on the port.



RADIUS Accounting

```
radius-server key plaintext aruba123

radius-server host 10.100.50.50 vrf mgmt
```

Server key and IP address

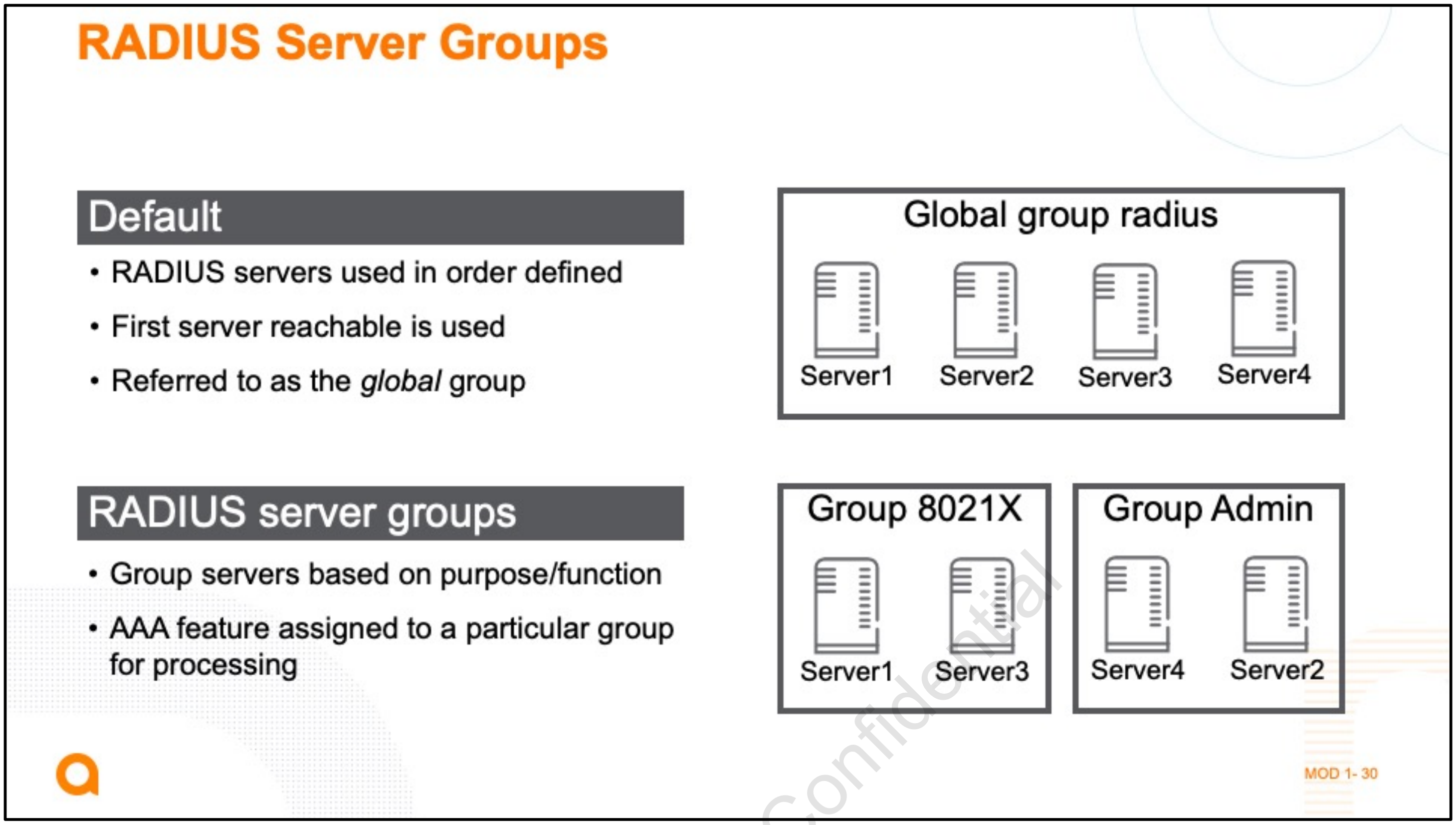
```
aaa accounting port-access start-stop group {radius | <GROUP-NAME>}
aaa accounting port-access start-stop interim <minutes> {radius | <GROUP-NAME>}
```

MOD 1- 29

Enable RADIUS for start-stop accounting as shown in the figure. Accounting will track when a user session begins and ends (start-stop).

For more detailed session tracking, you can configure the switch to send periodic accounting messages during the session, with interim option, as shown. The default interval is 60 minutes—in most cases you'll want to make this a smaller interval for more frequent updates to your statistics. Refer to your AOS-CX switch manual's list of network accounting attributes for a complete list of the RADIUS accounting attributes that it supports for 802.1X.

RADIUS accounting takes effect after a user successfully authenticates. The switch then creates a session for the user, defining the network access that the user is provided. If you selected start-stop, the switch sends a start accounting message to the RADIUS server when the session begins. With either stop-only or start-stop, the switch sends a message when the session ends (because the user disconnects or the session times out, for example). These messages include several standard RADIUS attributes that describe the user's identity (User-Name), location (NAS-IP-Address and NAS-Port), and bandwidth consumption (Acct-Output-Packets and Acct-Input-Packets), as well as other information.



RADIUS Server Groups

Default

- RADIUS servers used in order defined
- First server reachable is used
- Referred to as the *global* group

RADIUS server groups

- Group servers based on purpose/function
- AAA feature assigned to a particular group for processing

Global group radius

Server1 Server2 Server3 Server4

Group 8021X

Server1 Server3

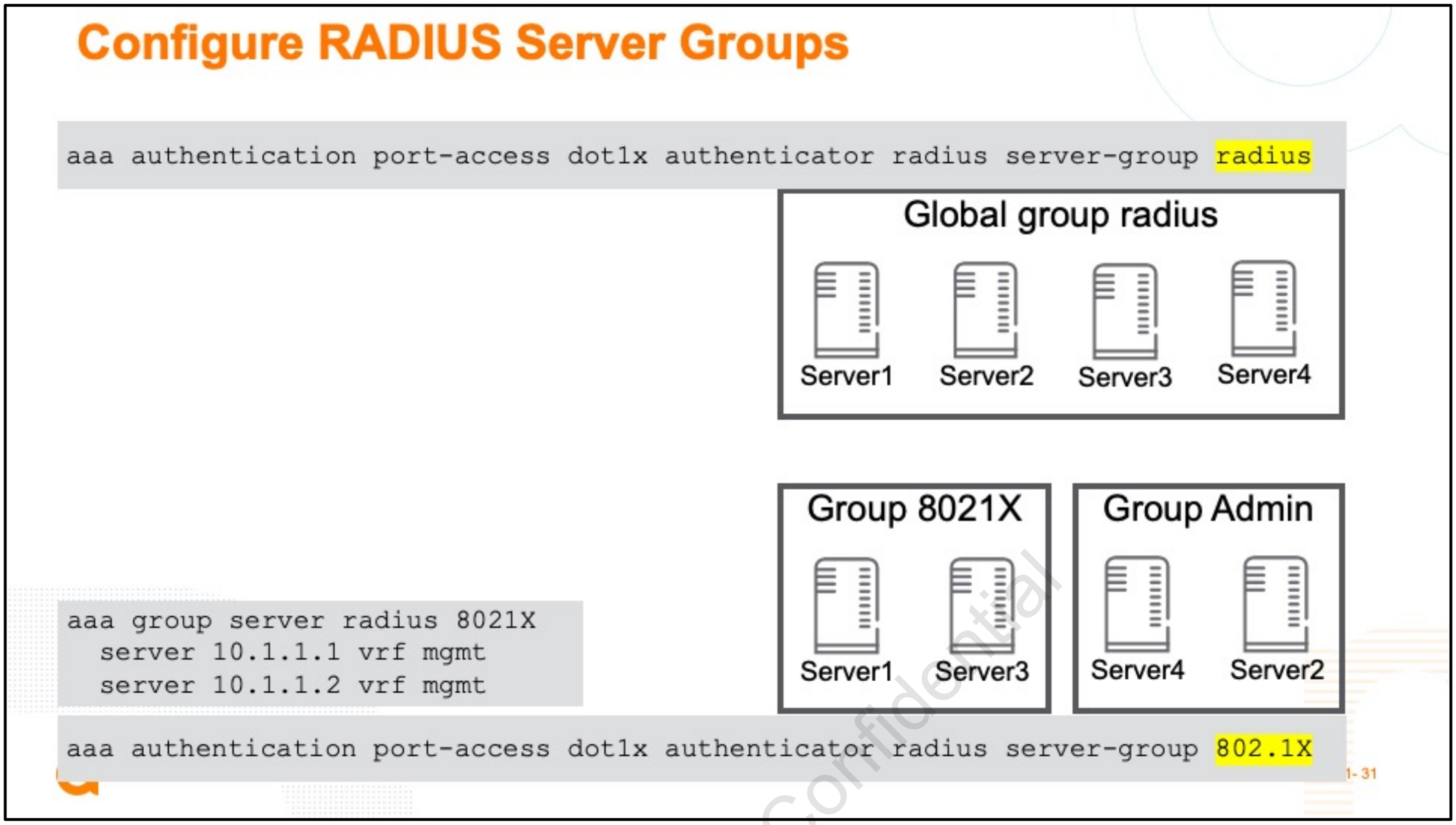
Group Admin

Server4 Server2

MOD 1- 30

Companies deploy multiple RADIUS servers to provide redundancy. The servers that you add are automatically added to the global RADIUS server group, which is called radius, and referred to as the global group. The switch uses this group, by default, for all types of authentication for which you have enabled RADIUS. Servers are used in the order they are defined.

You might want to use only a subset of servers for a particular task, instead of all the globally defined ones. You can do this by creating RADIUS groups – maximum 28 groups per switch. As shown in the figure, you might create a group of servers for end user authentication using 802.1X, and another group used for administrative access to the network devices' CLI and GUI.



Configure RADIUS Server Groups

```
aaa authentication port-access dot1x authenticator radius server-group radius
```

```
aaa group server radius 8021X
  server 10.1.1.1 vrf mgmt
  server 10.1.1.2 vrf mgmt
```

```
aaa authentication port-access dot1x authenticator radius server-group 802.1X
```

1-31

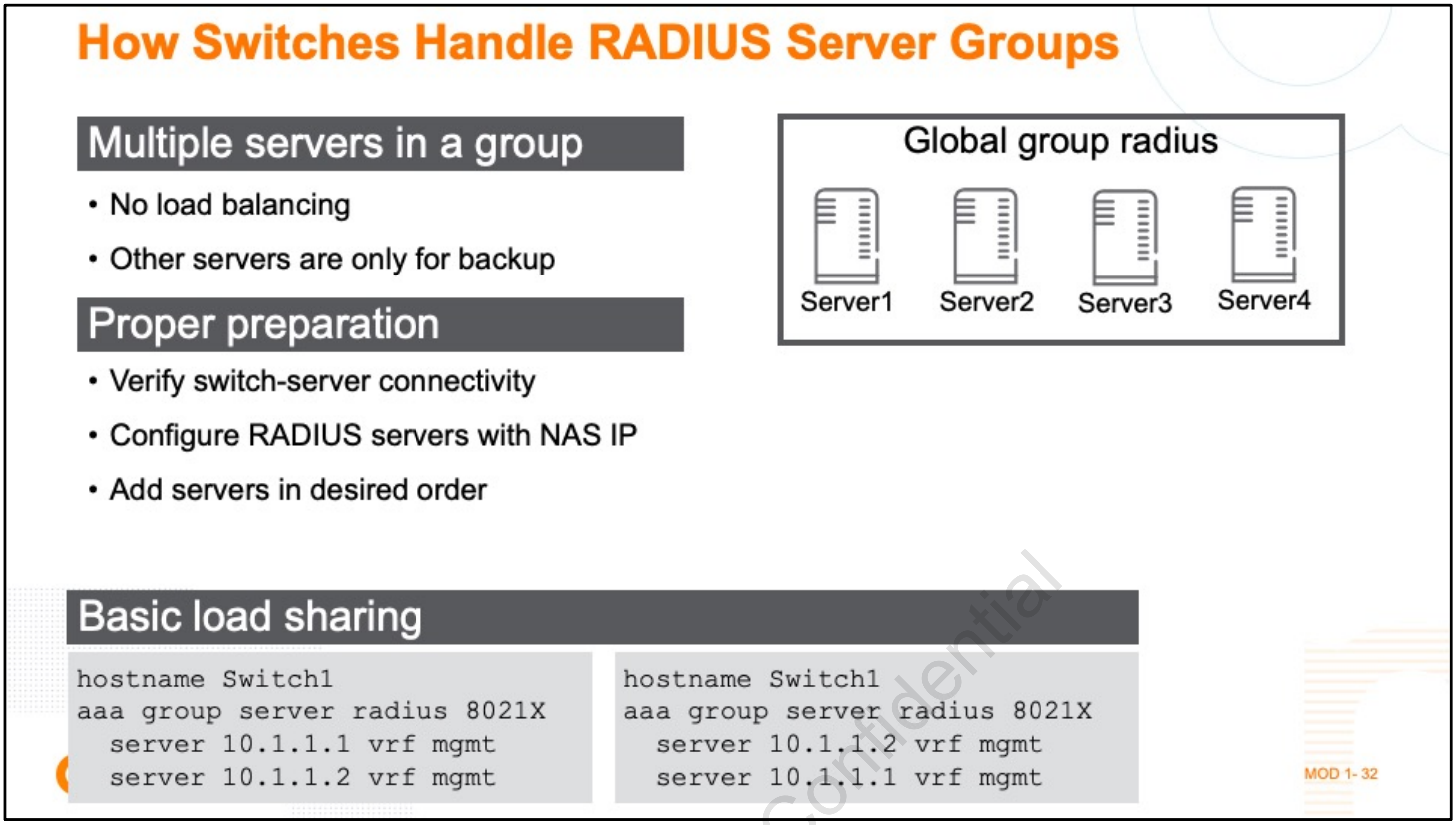
The top example in the figure shows how to configure RADIUS with 802.1X using the global default. The switch uses the global radius group to authenticate users who connect to ports with 802.1X enabled on them.

The bottom example shows how to configure RADIUS groups. First you define the group and its member servers. Then you indicate that the 802.1X process should use this group to authenticate users.

See which RADIUS servers are part of any group with the command show radius-server.

See which RADIUS servers are port of the radius group with the command show aaa server-group radius.

You cannot edit the radius server group. All globally defined RADIUS servers belong to it; however, you can create specific RADIUS server groups and include only the servers you want for the RADIUS service in question, like 802.1X or MAC authentication, for example.



How Switches Handle RADIUS Server Groups

Multiple servers in a group

- No load balancing
- Other servers are only for backup

Proper preparation

- Verify switch-server connectivity
- Configure RADIUS servers with NAS IP
- Add servers in desired order

Global group radius
Server1
Server2
Server3
Server4

Basic load sharing

```
hostname Switch1
aaa group server radius 8021X
  server 10.1.1.1 vrf mgmt
  server 10.1.1.2 vrf mgmt
```

```
hostname Switch1
aaa group server radius 8021X
  server 10.1.1.2 vrf mgmt
  server 10.1.1.1 vrf mgmt
```

MOD 1- 32

When a RADIUS server group has multiple servers, the switch does not load balance requests to different servers. Instead it sends all requests to the first server in the group that it successfully accesses. It uses the other servers as backups in case it cannot reach the first server.

For this reason, you should prepare before adding RADIUS servers to your switch and enable 802.1X. Ensure that the switch NAS has connectivity to the RADIUS servers, and the RADIUS servers are configured to accept messages from it. You should then add the RADIUS servers in the order in which you want the switch to use them.

For rudimentary load sharing, specify different servers first on different switches, as shown in the figure. So, the default timeout is 5 seconds and the default retransmit setting is 3. After 20 seconds, the switch determines that the RADIUS server is not available. The switch then tries to reach the next RADIUS server in the list. If the switch cannot reach that RADIUS server either, after 20 seconds, it tries the next server on the list. If the switch cannot reach any of the servers, it tries the first server in the list again.

However, you can configure the switch to wait a set period after the last failed retransmit to a particular server before attempting to communicate with that server again. You might do this to prevent a switch from endlessly attempting to send RADIUS packets during a network outage. This period is called the dead time.

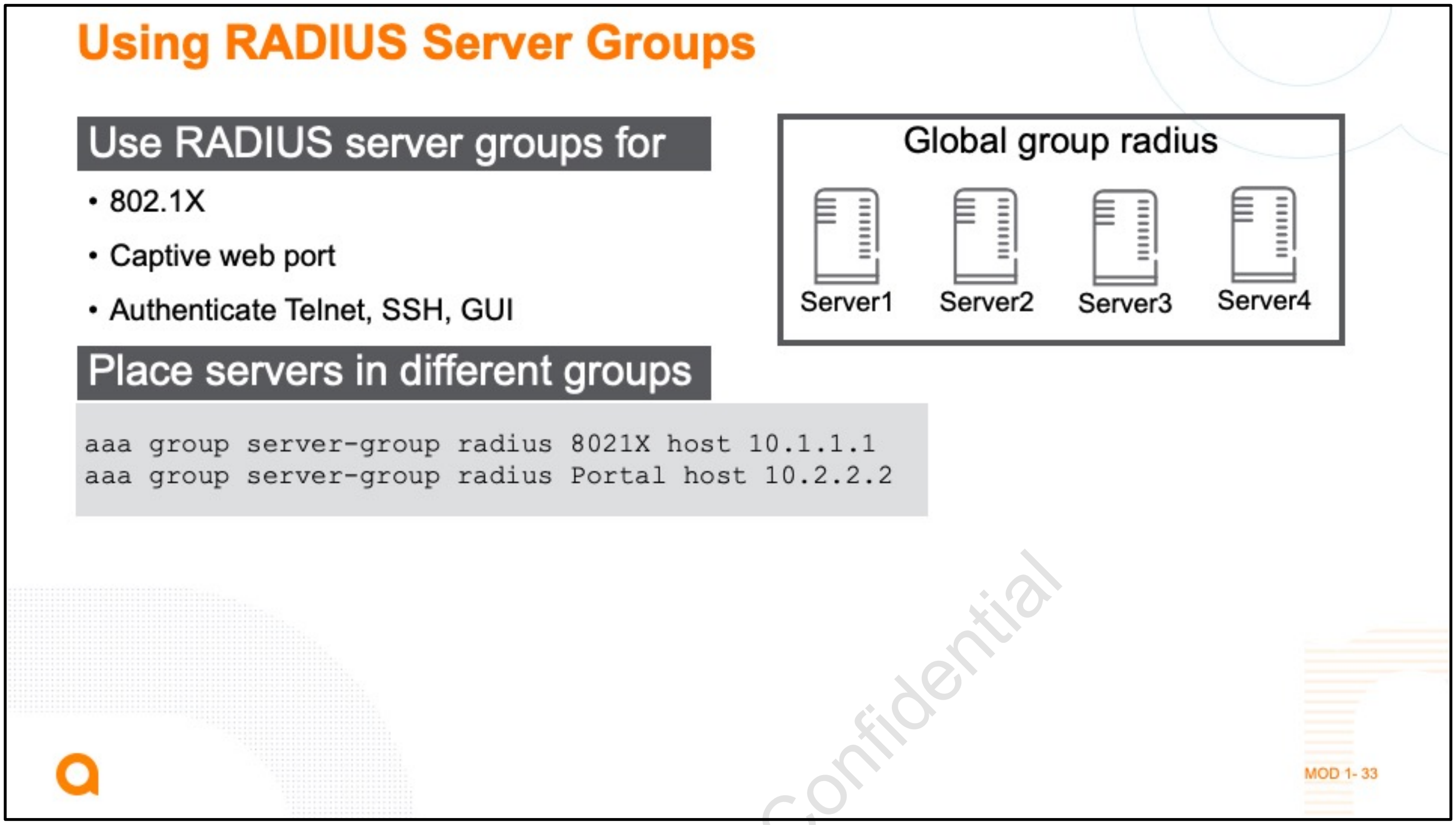
These are global settings that apply to all RADIUS servers and groups. You can adjust the timeout (in seconds—5 is the default) and the number of retransmits (1 is the default), and the dead time (in minutes) with these commands:

```
Switch(config)# radius-server timeout <1-60>
Switch(config)# radius-server retries <0-5> fail-through
```

By default, if a RADIUS server is reachable and the user fails authentication, the switch uses this result. However, if the user is a visitor from a different location, like a neighboring campus, and their credentials are not stored on the local campus server, but the remote campus server, the user would be denied access by default.

You can enable the fail-through feature, so that if the first RADIUS server is reachable and returns a “failed” result, the switch can proceed and try other RADIUS servers in the list or group. This feature is disabled by default, but can be enabled with this command:

```
Switch(config)# aaa authentication allow-fail-through
```



Using RADIUS Server Groups

Use RADIUS server groups for

- 802.1X
- Captive web port
- Authenticate Telnet, SSH, GUI

Place servers in different groups

```
aaa group server-group radius 8021X host 10.1.1.1
aaa group server-group radius Portal host 10.2.2.2
```

Global group radius
Server1
Server2
Server3
Server4

MOD 1- 33

Other RADIUS server groups

This course module focuses on 802.1X authentication. However, as you will learn in other modules, AOS-CX switches can use RADIUS with other types of authentication such as captive portal and the authentication of Telnet, SSH, or web UI managers.

To use different RADIUS servers for different types of authentication, you must place those servers in different RADIUS server groups, as shown in the figure. Use the no form of the command to delete the server from the group. Each group can have up to three RADIUS servers.

A RADIUS server can be part of more than one group. If, for example, you add a server that is part of the radius server group to a new group, the server remains part of the radius group, as well. Therefore, you should define the servers that you want to use for the global group before adding the servers for the other groups. Otherwise, the servers that you intend for other groups will take up the slots in the global radius group.

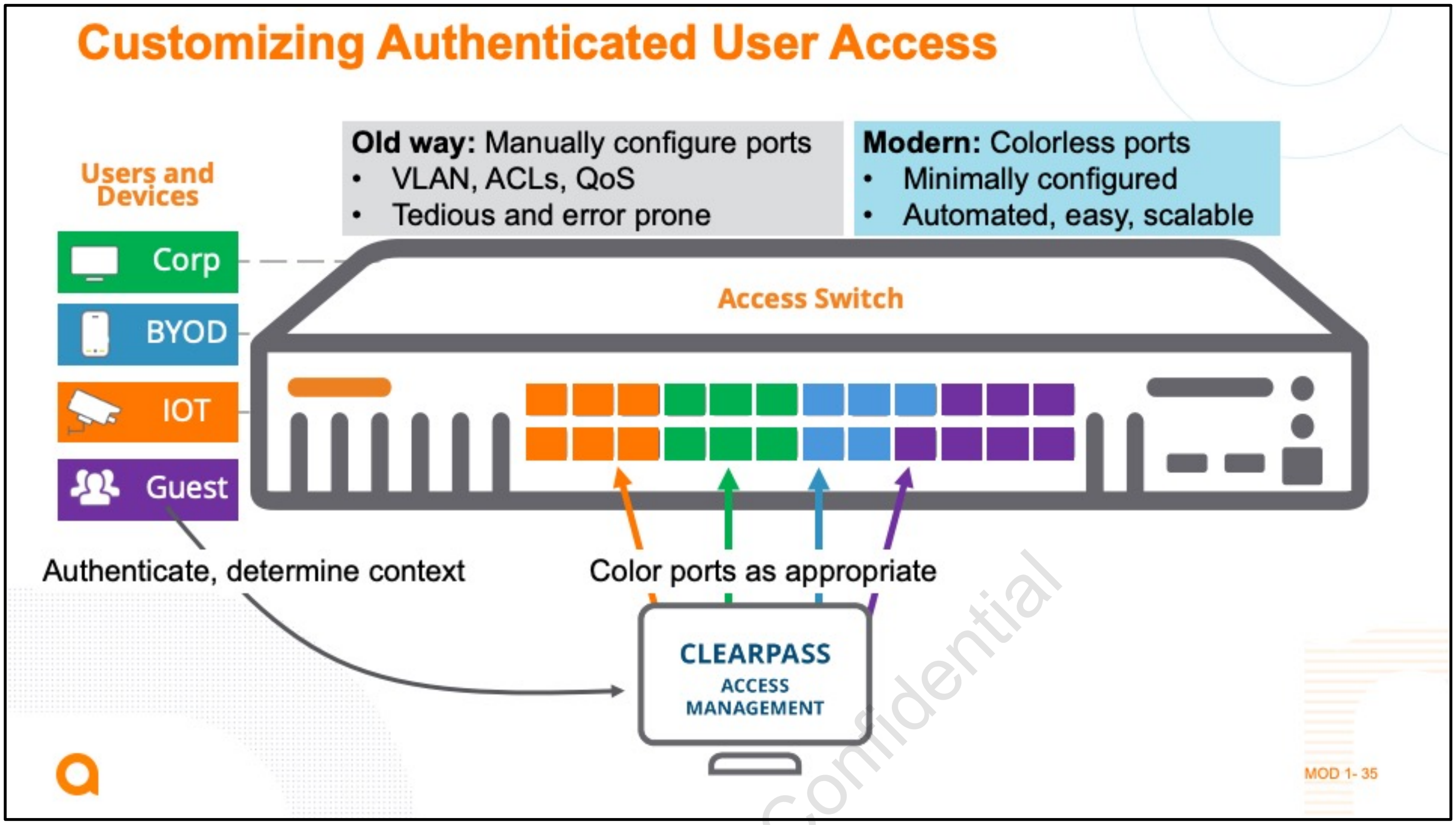
Remember, to use the RADIUS server group, you must choose the type of authentication to which it applies:

```
Switch(config)# aaa authentication port-access eap-radius server-group <group name>
```

RADIUS Attributes

MOD 1- 34

You will now examine some of the more advanced scenarios for 802.1X authentication.



Customizing Authenticated User Access

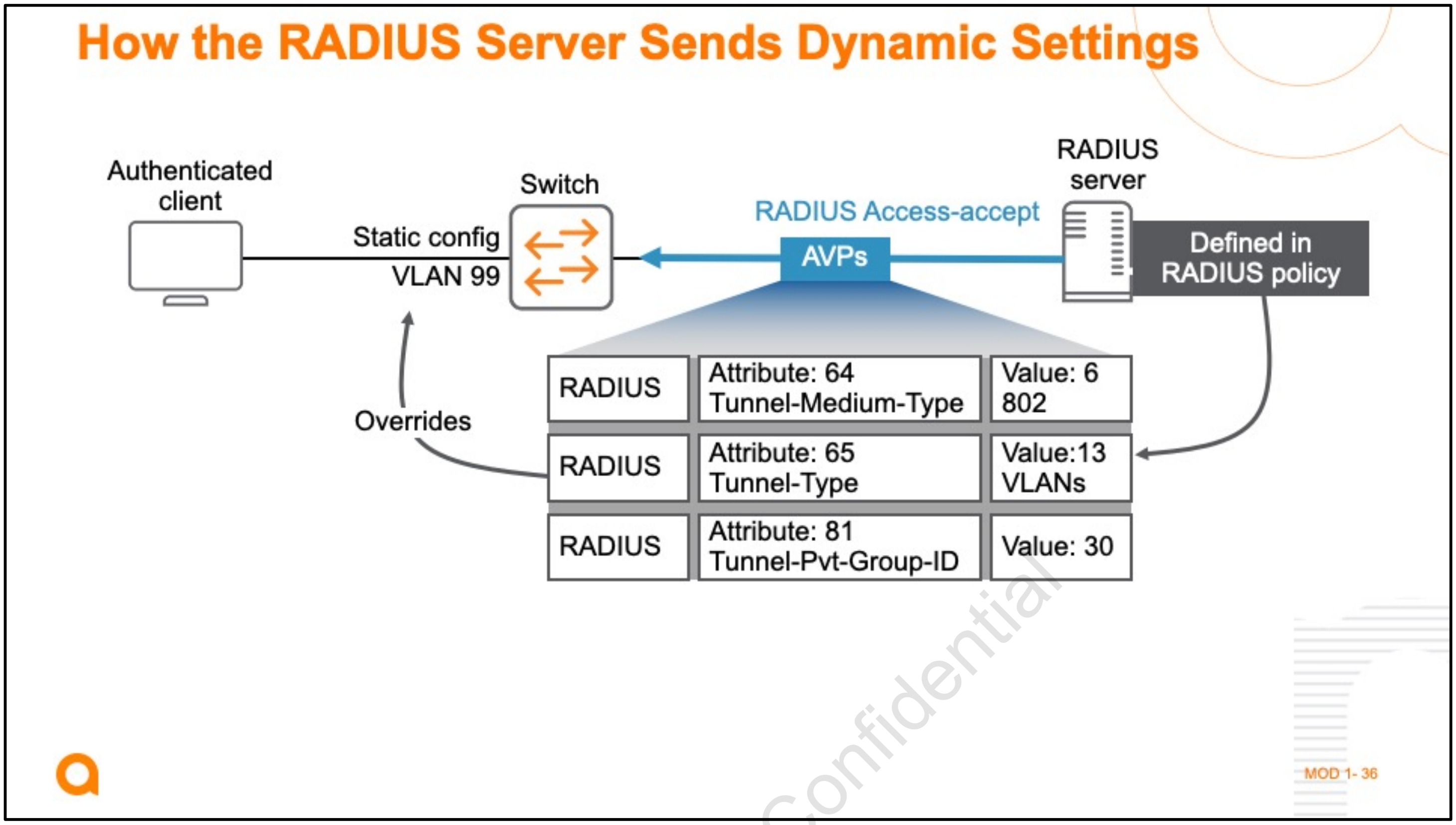
Old way: Manually configure ports
- VLAN, ACLs, QoS
- Tedious and error prone

Modern: Colorless ports
- Minimally configured
- Automated, easy, scalable

Beyond the security benefits, port-based 802.1X authentication helps you to provision the network edge more quickly. You learned about the old days, when you had to configure each port with the appropriate settings for the intended user – tedious and error prone.

With 802.1X, the RADIUS server and NAS switch automatically determine user identity as well as other context characteristics – who, what, how, where, when. A RADIUS server policy can use this information to automatically configure appropriate settings – VLAN, ACLs, QoS/CoS, rate limits, and more. The network edge becomes more flexible and adaptable.

Because the switch applies these settings to a user session automatically, as dictated by the RADIUS server, these settings are sometimes called RADIUS settings or dynamic settings. For example, a dynamic VLAN assignment is the VLAN assigned to a user session by the RADIUS server when the user authenticates to the network. A dynamic ACL is the (port-based) ACL applied to a user session by the RADIUS server when the user authenticates, and so on.



How the RADIUS Server Sends Dynamic Settings

RADIUS	Attribute: 64 Tunnel-Medium-Type	Value: 6 802
RADIUS	Attribute: 65 Tunnel-Type	Value:13 VLANs
RADIUS	Attribute: 81 Tunnel-Pvt-Group-ID	Value: 30

MOD 1- 36

You define dynamic policy settings on a RADIUS server like ClearPass. These settings apply to certain users based on those identity and context. The server communicates these settings to the NAS switch in Attribute-Value Pairs (AVPs) within the RADIUS-Accept message. They are called Attribute-Value Pairs because each data item has an attribute type and an associated value. The figure shows three AVPs:

Attribute Type 64 defines the type of medium over which this communication occurs, with a value of 6, which means an 802-type “tunnel”

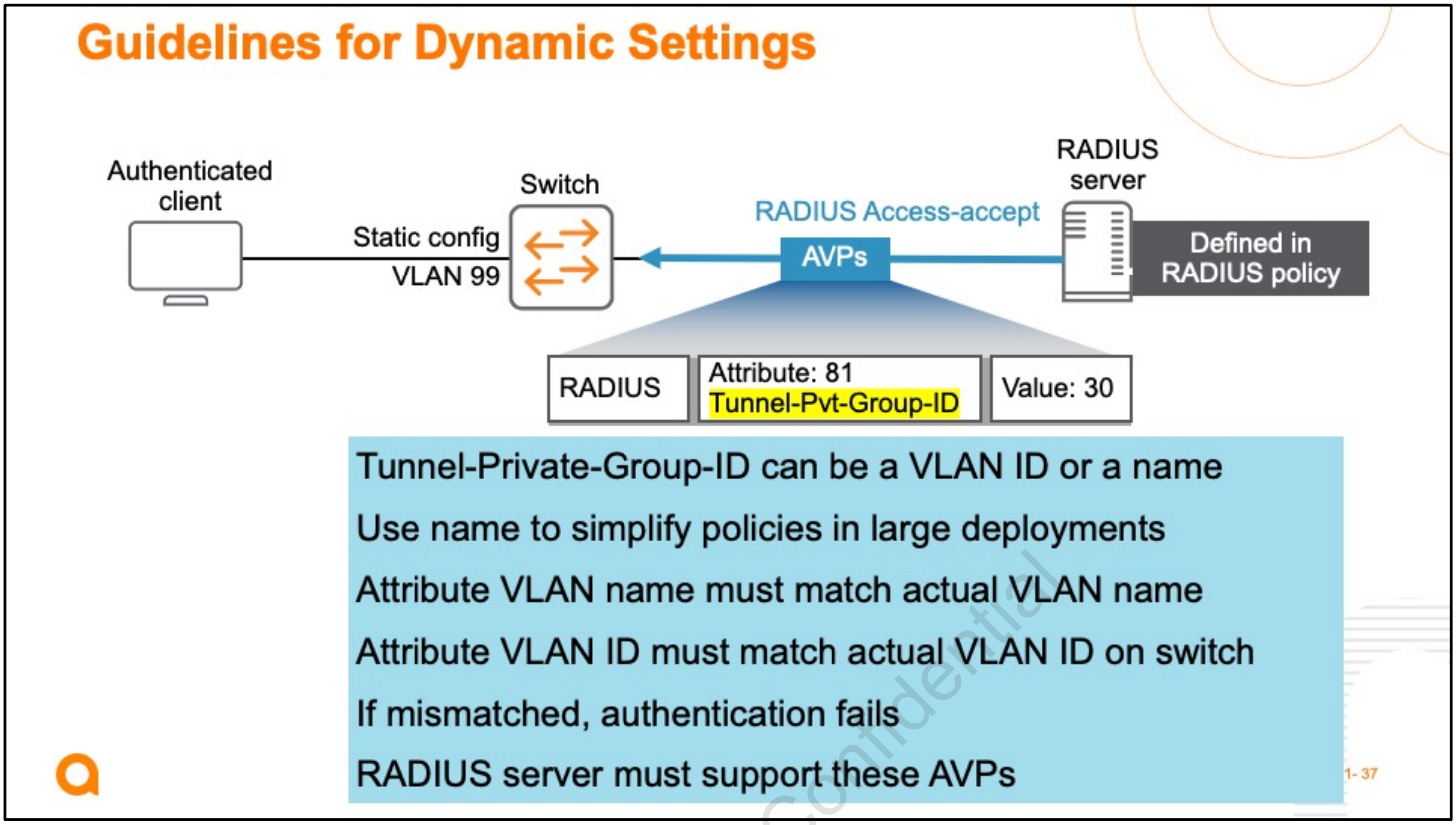
Attribute Type 65 defines the tunnel type, which is set to 13. This is used to indicate that we intend to assign a VLAN ID.

Attribute Type 81 defines the actual VLAN ID – in this case, VLAN 30.

Thus, the purpose of this set of AVPs sent from RADIUS server to switch is to override the port’s statically configured VLAN 99 with VLAN 30.

There are standard AVPs defined by the IETF in an RFC, and there are custom attributes that are defined specifically by a vendor like Aruba. These are called Vendor-Specific Attributes (VSA). There are also extended RADIUS attributes.

When possible, you should use standard RADIUS attributes because they are supported across many vendors. Depending on the software version, AOS-CX switches might support multiple attributes for some settings. Generally, earlier software used vendor specific attributes while the more up-to-date software added standard RADIUS attributes for applying the same setting. Aruba recommends that you typically use the standard RADIUS setting, and only use the vendor specific one if you need to accommodate a legacy solution.



Guidelines for Dynamic Settings

Authenticated client

Static config
VLAN 99

Switch

RADIUS Access-accept

AVPs

RADIUS server

Defined in RADIUS policy

RADIUS	Attribute: 81 Tunnel-Pvt-Group-ID	Value: 30

Tunnel-Private-Group-ID can be a VLAN ID or a name

Use name to simplify policies in large deployments

Attribute VLAN name must match actual VLAN name

Attribute VLAN ID must match actual VLAN ID on switch

If mismatched, authentication fails

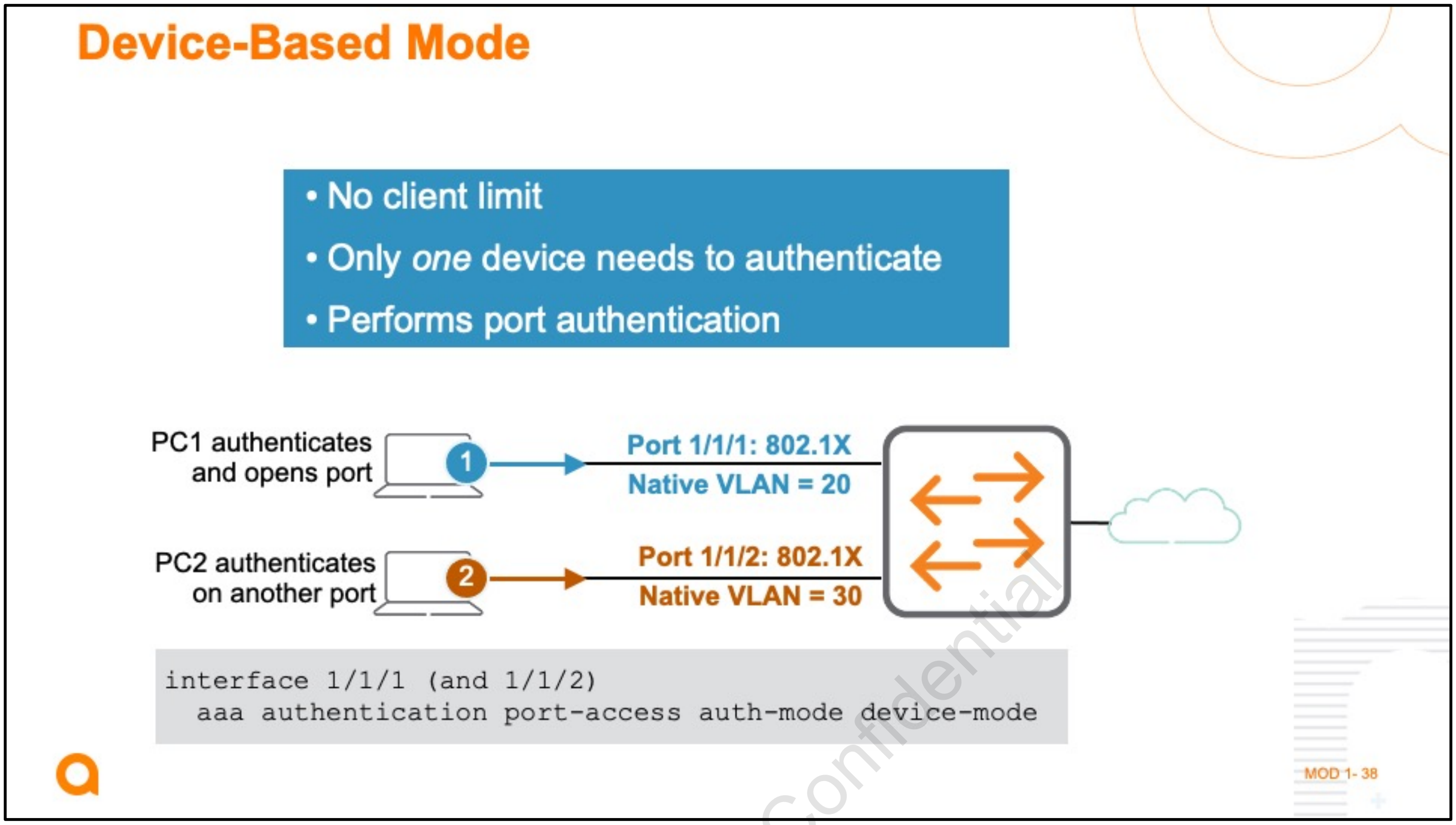
RADIUS server must support these AVPs

1- 37

Keep a few guidelines in mind when coordinating the value for a dynamic VLAN with ClearPass administrators. The Tunnel-Private-Group-ID can be a VLAN ID or a name. You might use the name option to simplify policies in a large or multi-site network. For example, switches in one building might name VLAN 10 “Students” while switches in another building name VLAN 11 the same name. The RADIUS server could use a single policy to assign students to the “Students” VLAN with the Egress-VLAN-Name AVP. The switch through which the individual student authenticated would then assign the student’s untagged traffic (native VLAN) to the correct VLAN ID associated with the “Students” name on that switch.

Be very careful, though, that the name in the attribute matches the name of a VLAN on the switch; names are case sensitive. Similarly, a VLAN ID must match a VLAN on the switch. If the RADIUS server sends a VLAN name or ID that does not exist on the switch, the switch does not apply the port’s static VLAN to the user session. Instead, the authentication fails—much as it would if the user entered an incorrect password.

Make sure that the RADIUS server supports these AVPs. Also make sure that the administrators know to define the vendor-specific AVPs, when required.



Device-Based Mode

- No client limit
- Only *one* device needs to authenticate
- Performs port authentication

```
interface 1/1/1 (and 1/1/2)
  aaa authentication port-access auth-mode device-mode
```

MOD 1- 38

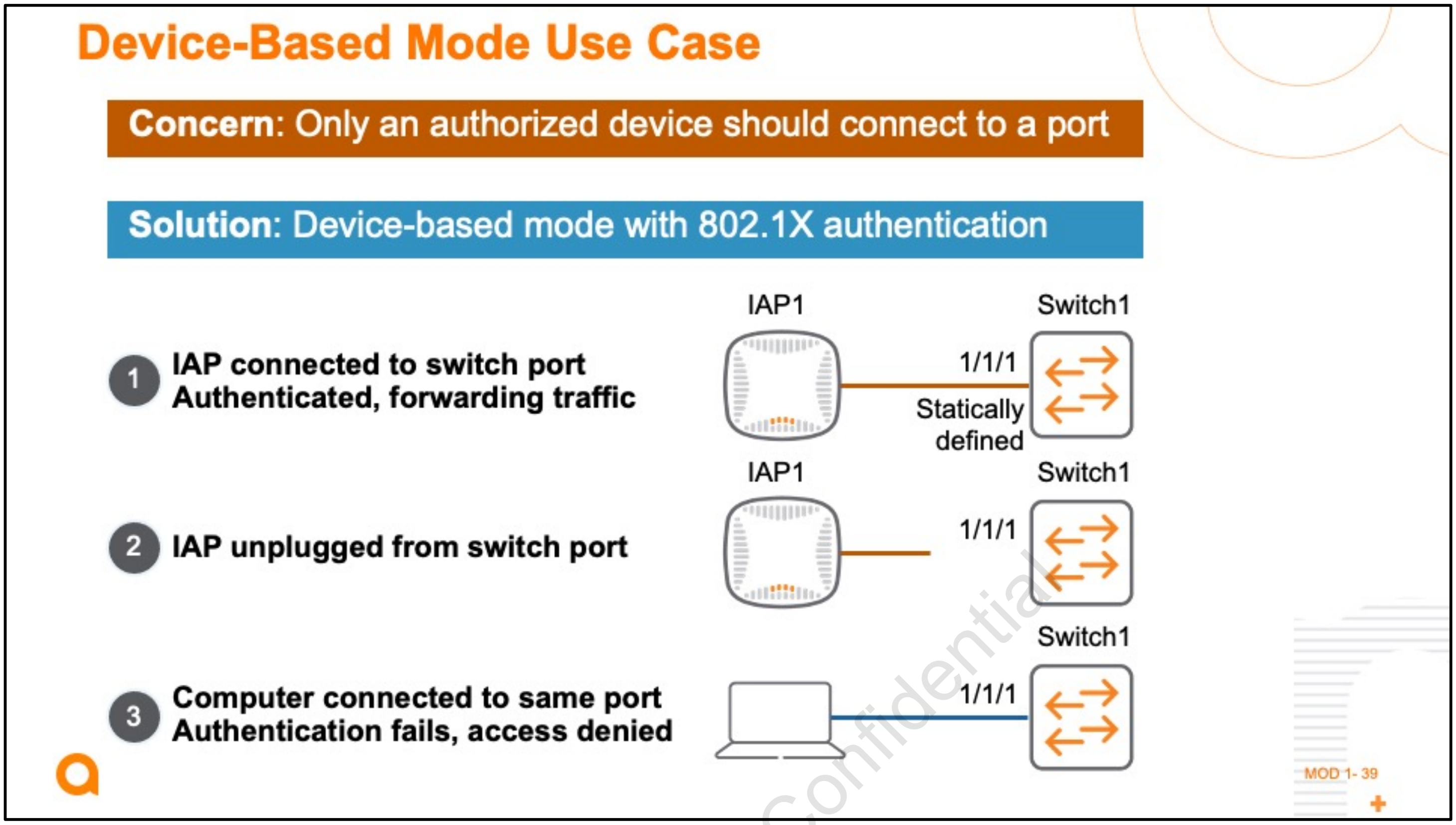
A switch applies AVP settings based on whether the port operates in 802.1X port-based or user-based mode. The mode also determines how the port establishes authentication sessions.

With device-based/port-based control, the switch establishes a single session for all traffic sent and received on the port. If any one user authenticates, the switch permits all traffic without referring to the source or destination MAC address. Untagged traffic (native VLAN), regardless of source MAC, is forwarded in the native VLAN assigned to the port (either statically or through RADIUS, as you have learned).

Often, only a single device connects to the switch port, so all traffic does originate from a single source MAC address. The example above illustrates this situation, where Computer1 and Computer2 are connected to two different switch ports and are assigned different VLANs based on their 802.1X credentials.

Configure the port mode for device mode as shown in the figure.

Note: Client mode is the default mode of the port.



Device-Based Mode Use Case

Concern: Only an authorized device should connect to a port

Solution: Device-based mode with 802.1X authentication

1. **IAP connected to switch port Authenticated, forwarding traffic**
2. **IAP unplugged from switch port**
3. **Computer connected to same port Authentication fails, access denied**

The concern with some customers is that the access switch might be in an unsecured environment.

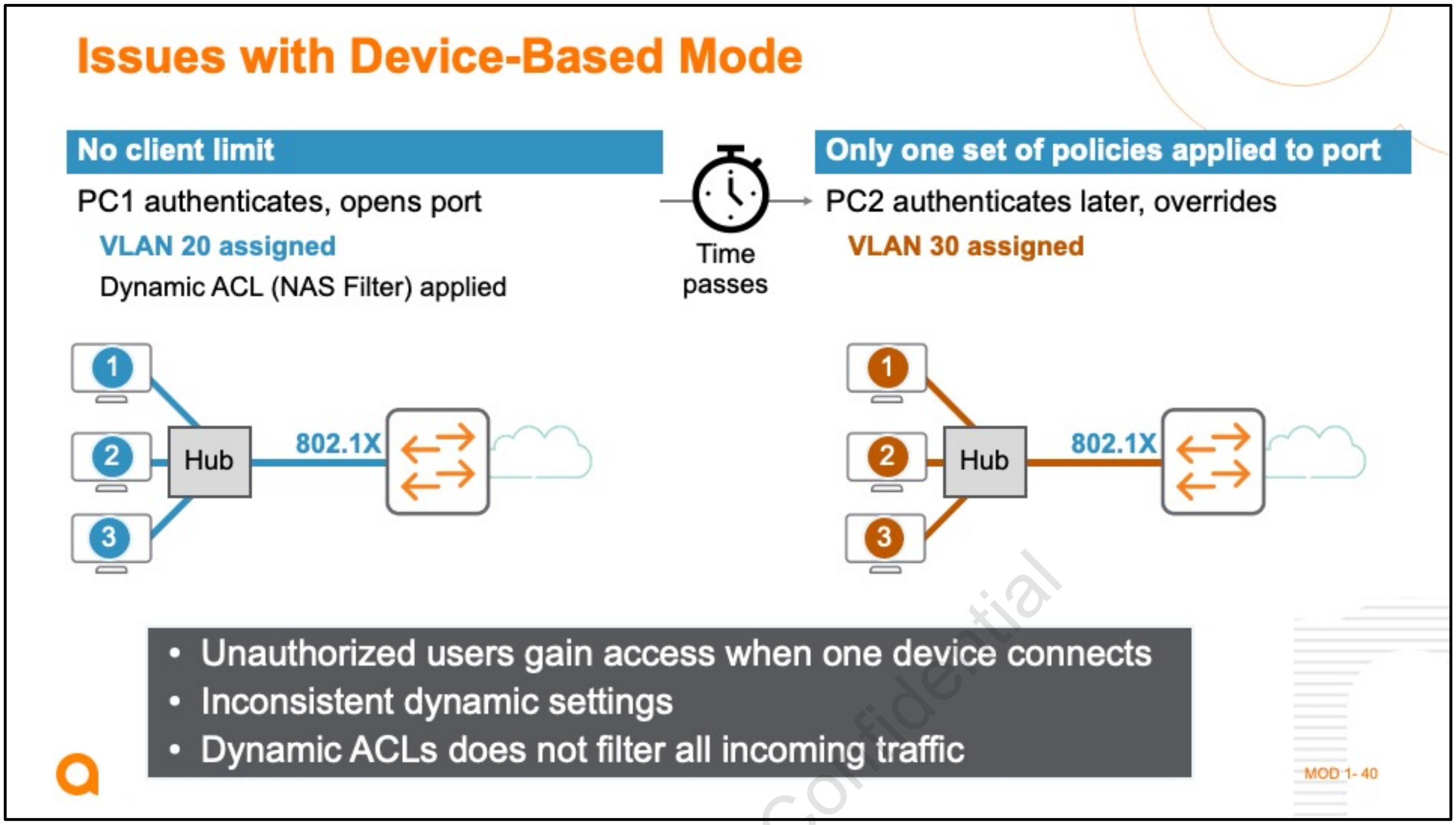
| Suppose a company statically defines switch port policy settings for Instant APs (IAPs).

| Someone unplugs an IAP from statically-defined port 1/1/1 and connects an unauthorized device to that port. The unauthorized device would have the same policy settings applied as the IAP.

If this was a corporate user, they would not have the correct VLAN settings, access would probably be denied. If this was a bad actor, they might be able to gain unauthorized access to corporate resources, depending on how the port is configured.

Device-based authentication mitigates this issue. If someone unplugged the IAP and plugged in an authorized device in this situation, the unauthorized device would be denied access to the network unless it successfully authenticated.

Note: Device-based mode is useful when you know that you'll only have one device connected to a port, you need to authenticate access on the port, and want to apply policies dynamically to the port based on authentication credentials (and possibly other information, like fingerprinting).



Issues with Device-Based Mode

No client limit

PC1 authenticates, opens port

VLAN 20 assigned

Dynamic ACL (NAS Filter) applied

Time passes

Only one set of policies applied to port

PC2 authenticates later, overrides

VLAN 30 assigned

- Unauthorized users gain access when one device connects
- Inconsistent dynamic settings
- Dynamic ACLs does not filter all incoming traffic

MOD 1- 40

An 802.1X-enabled interface might also connect to multiple devices. You could have an old Ethernet hub or unmanaged switch that does not know how to process EAPOL frames – so the switch or hum simply floods them. The figure shows how port-based mode interfaces handle this situation.

The interface has a single untagged (native) VLAN assignment. This could be a statically assigned VLAN, or a dynamic VLAN sent by the RADIUS server when the most recently authorized user authenticated. In the figure, user 1 authenticates and assigned to VLAN 20. User 2 and 3 have not authenticated, but they can also send traffic in VLAN 20. Remember that with Device-based mode, when one device authenticates, all are authenticated.

Understand that AOS-CX switches automatically bind a dynamic ACL to a user’s MAC address, even when 802.1X operates in port-based mode on the interface. Therefore, unauthenticated devices can receive access, and their traffic is not filtered by the dynamic ACL. Any static port ACL (PACL) on the port applies to both the authenticated client and the unauthenticated ones—as do in VACLs or RACLs that would normally filter the devices’ traffic.

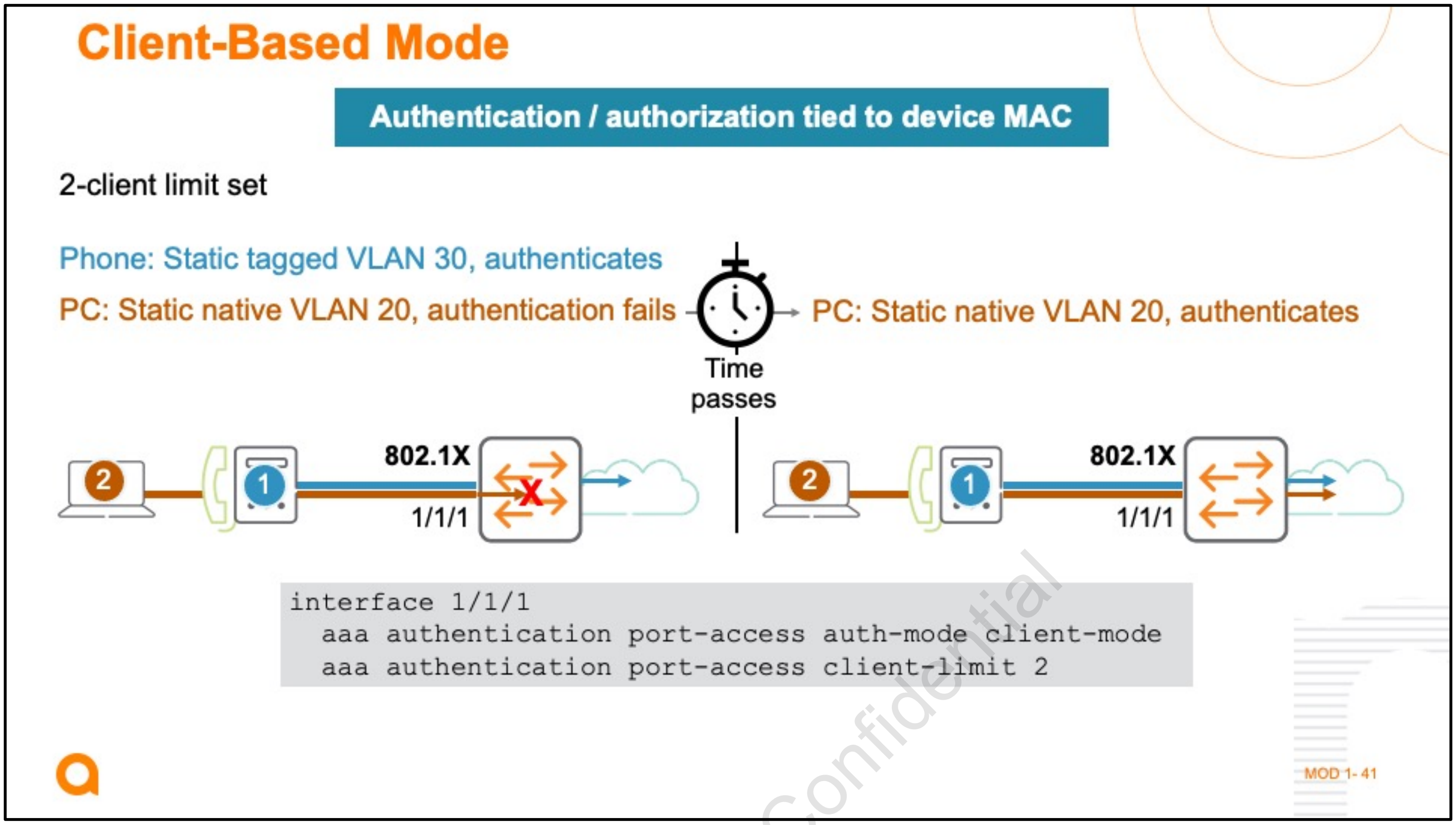
Suppose that some time later, PC2 authenticates. The RADIUS server sends new dynamic settings, which override those from the first device. This can cause connectivity issues, as the devices are moved to a new VLAN.

If only one device connects to the port at once, port-based mode can work well. But if multiple devices connect to the same port, device (port) based control introduces these potential issues:

Unauthorized users gain access by “piggybacking” on the same connection as an authorized device

Inconsistent dynamic settings

Dynamic ACLs does not filter all incoming traffic



Client-Based Mode

Authentication / authorization tied to device MAC

2-client limit set

Phone: Static tagged VLAN 30, authenticates

PC: Static native VLAN 20, authentication fails

PC: Static native VLAN 20, authenticates

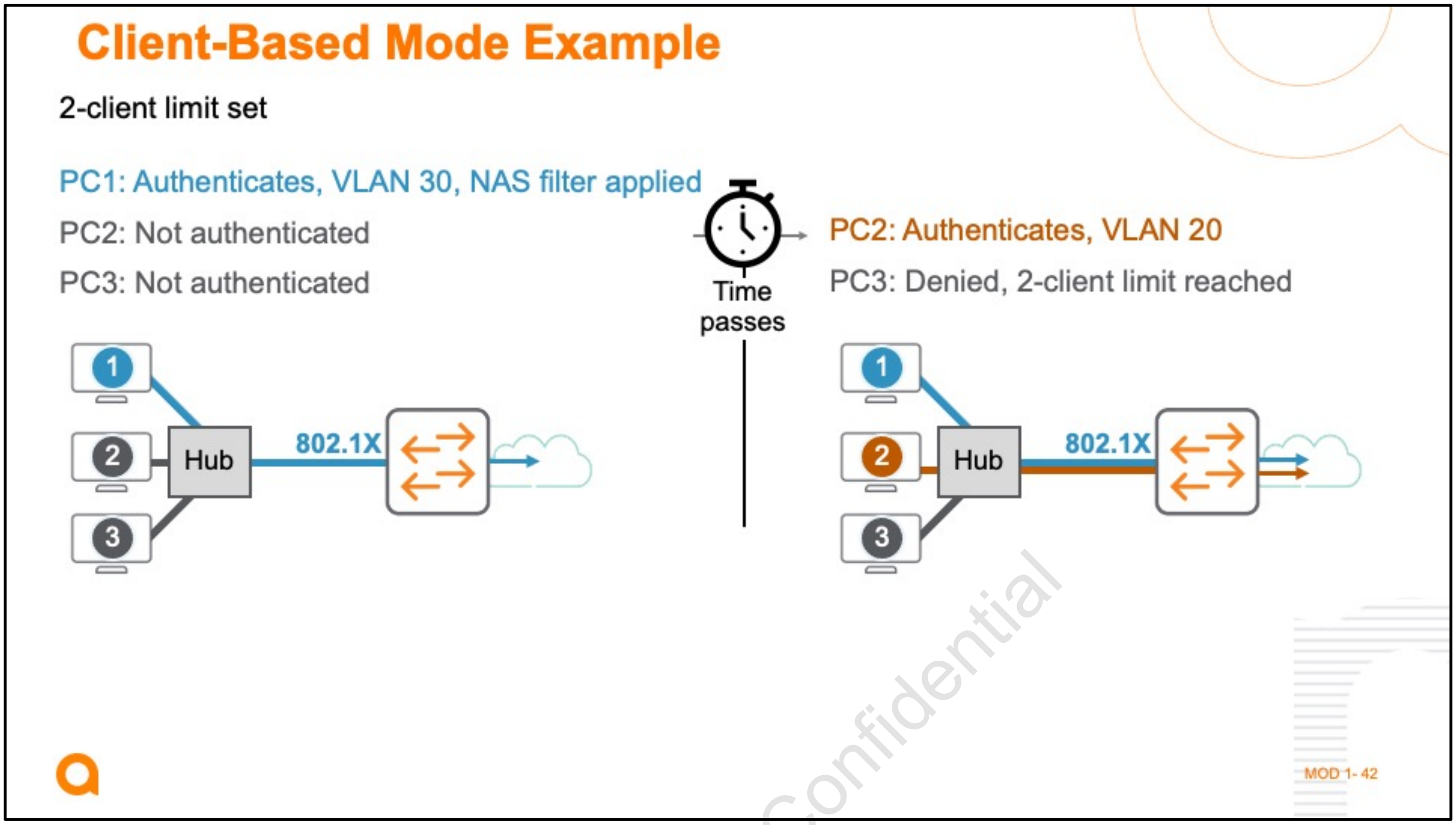
```
interface 1/1/1
  aaa authentication port-access auth-mode client-mode
  aaa authentication port-access client-limit 2
```

Client-based mode is recommended for scenarios in which you want the switch port itself to authenticate and control multiple clients. This is recommended when the RADIUS server might implement different dynamic settings for different clients. Often, you should also enable user-based control even when the port should only connect to a single device to prevent unauthorized devices from “piggybacking” on the port.

In user-based mode the port authenticates each downstream device, as identified by source MAC address. You enable this mode when you set the client limit on the port to any value including 1. Client mode is the default, with a default client limit of 1. See the configuration example in the figure.

The port creates a separate session for each device, and each session has its own authorized status. In other words, the switch permits ingress traffic with a particular source MAC address only if that MAC address’s session is authorized. Conversely, it permits egress traffic based on the destination MAC address’s authorization status. The limit value indicates how many devices can be authorized.

After you set the limit, the port begins tracking MAC addresses and defines the authorization status and settings for each separately. For example, in the scenario with the computer and VoIP phone, the switch port sends an EAP Request/Identity to each separate MAC address detected on the port. If the VoIP phone authenticates successfully, but the computer fails, the computer traffic is blocked.



Client-Based Mode Example

2-client limit set

PC1: Authenticates, VLAN 30, NAS filter applied

PC2: Not authenticated

PC3: Not authenticated

Time passes

PC2: Authenticates, VLAN 20

PC3: Denied, 2-client limit reached

MOD 1- 42

Compare this client-based example to the similar, previously shown example for device-based mode. Again, the client limit is set to two.

Initially, only PC1 is authenticated, so the switch drops traffic from PC2 and PC3. Then PC2 authenticates, so the switch forwards PC2 traffic. Now the limit is reached, so user 3 cannot authenticate until another user disconnects.

The examples above also illustrate how client-based mode allows each session to have its own dynamic attributes, such as VLAN and ACL. The dynamic VLAN and dynamic NAS filter only applies to PC1. The switch forwards any untagged, Native VLAN traffic sent by the user device in the dynamic VLAN. And when it needs to forward traffic to the device, it forwards the traffic without a tag. If the port supports multiple users, the users can have different native VLAN assignments. If any users do not receive a dynamic assignment, the static native VLAN applies to their traffic.

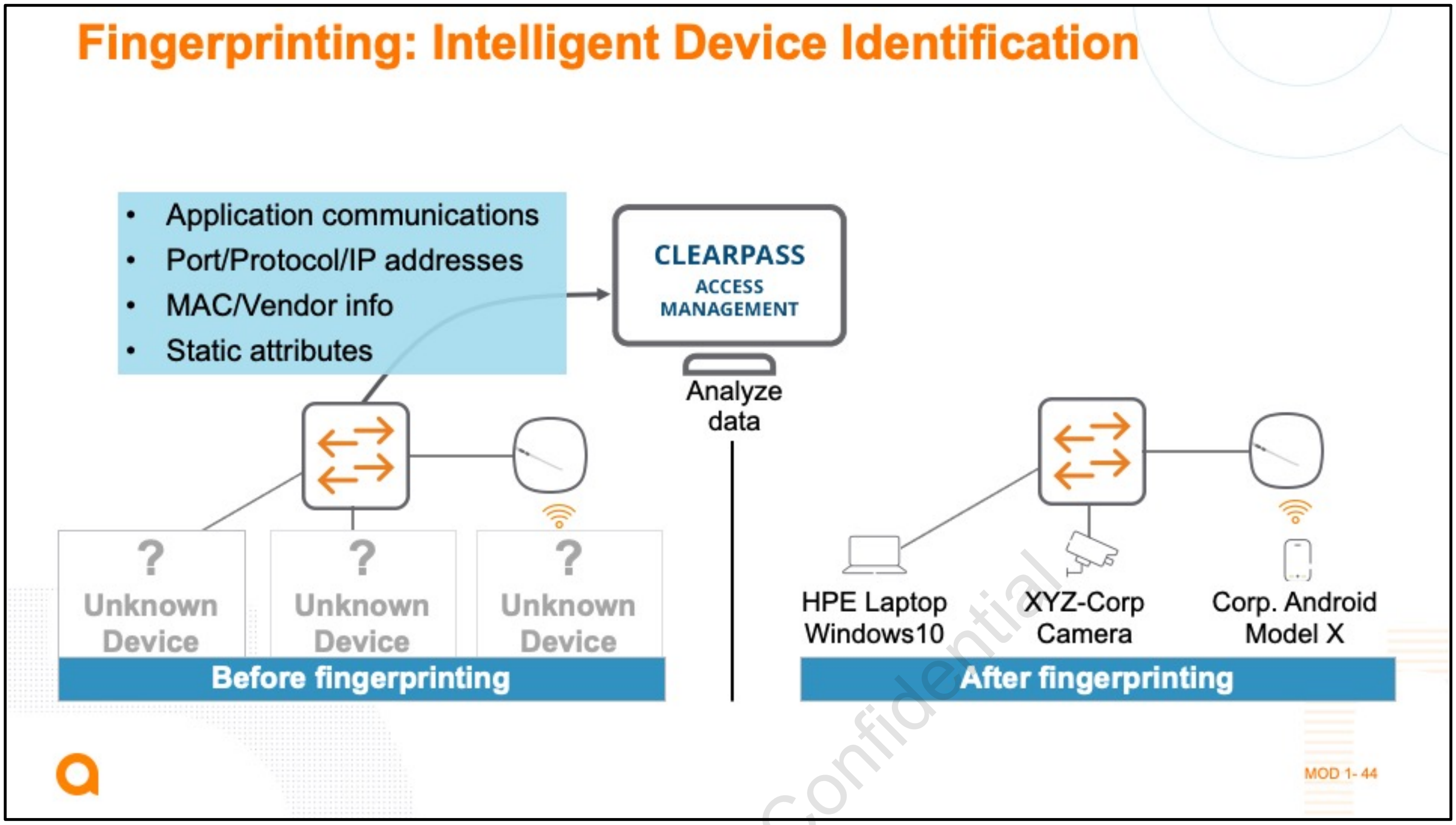
In the example, above both PC1 and PC2 are authenticated and assigned to different VLANs. The switch recognizes which user is sending traffic based on source MAC address and forwards the traffic in the correct VLAN.

Note that the switch must flood native VLAN traffic in all VLANs out of the port. This means that devices will receive broadcasts and multicasts for all untagged VLANs (native VLANs) applied to the port. Keep this in mind when you plan multi-client scenarios.

User Roles

MOD 1- 43

This section introduces a feature that was copied from the Mobility controllers, called user roles, and is implemented in AOS-CX switches.



Fingerprinting: Intelligent Device Identification

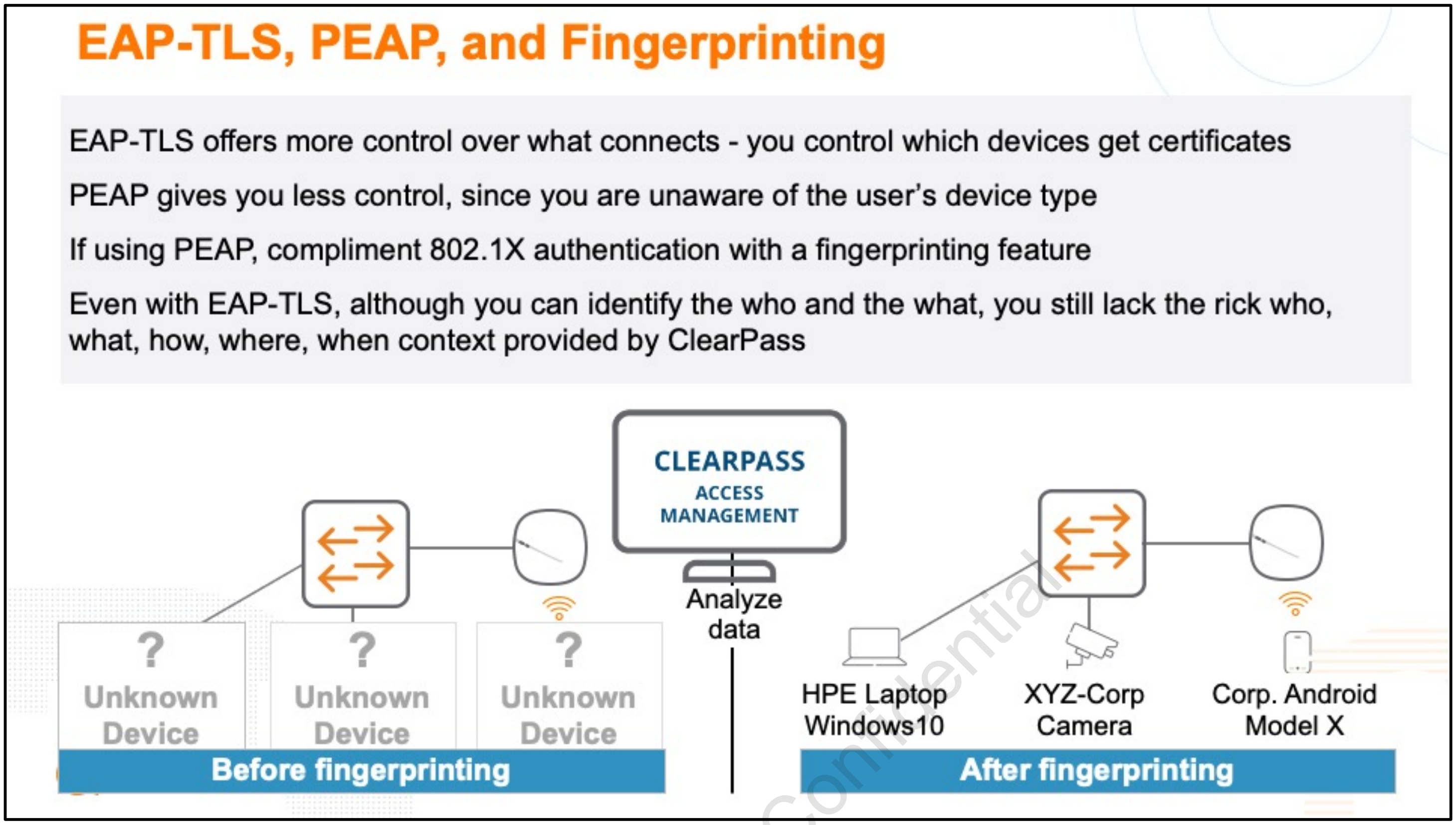
- Application communications
- Port/Protocol/IP addresses
- MAC/Vendor info
- Static attributes

MOD 1- 44

You learned the advantage of having colorless ports with minimal configuration. VLAN, QoS, and other policy information is dynamically configured based on what connects. Historically, this was done with RADIUS VSA attributes. However, Aruba supports a feature called user roles that can be used instead. User roles are supported on Aruba Mobility Controllers, AOS switches, and AOS-CX switches. User roles allow for a single process configuration that provides a consistent user experience – whether they connect to wired switch ports and wirelessly. You get simplified, centralized policies that reduce administrative overhead on improve scalability.

If using PEAP authentication, you know who the user is, but not what device they are using. ClearPass profiling capabilities uses fingerprinting to determine the device type and operating system.

Important: Using EAP-TLS with certificates gives you much more control over what is connecting to the network - you control which devices get certificates. Using PEAP for authentication gives you less control, since you don't know what device the user connects with. If using PEAP, you should compliment your 802.1X authentication with a fingerprinting feature like that supported by ClearPass. Even with EAP-TLS, although you can identify the who and the what, you still lack the rick who, what, how, where, when context provided by Clearpass.



EAP-TLS, PEAP, and Fingerprinting

EAP-TLS offers more control over what connects - you control which devices get certificates

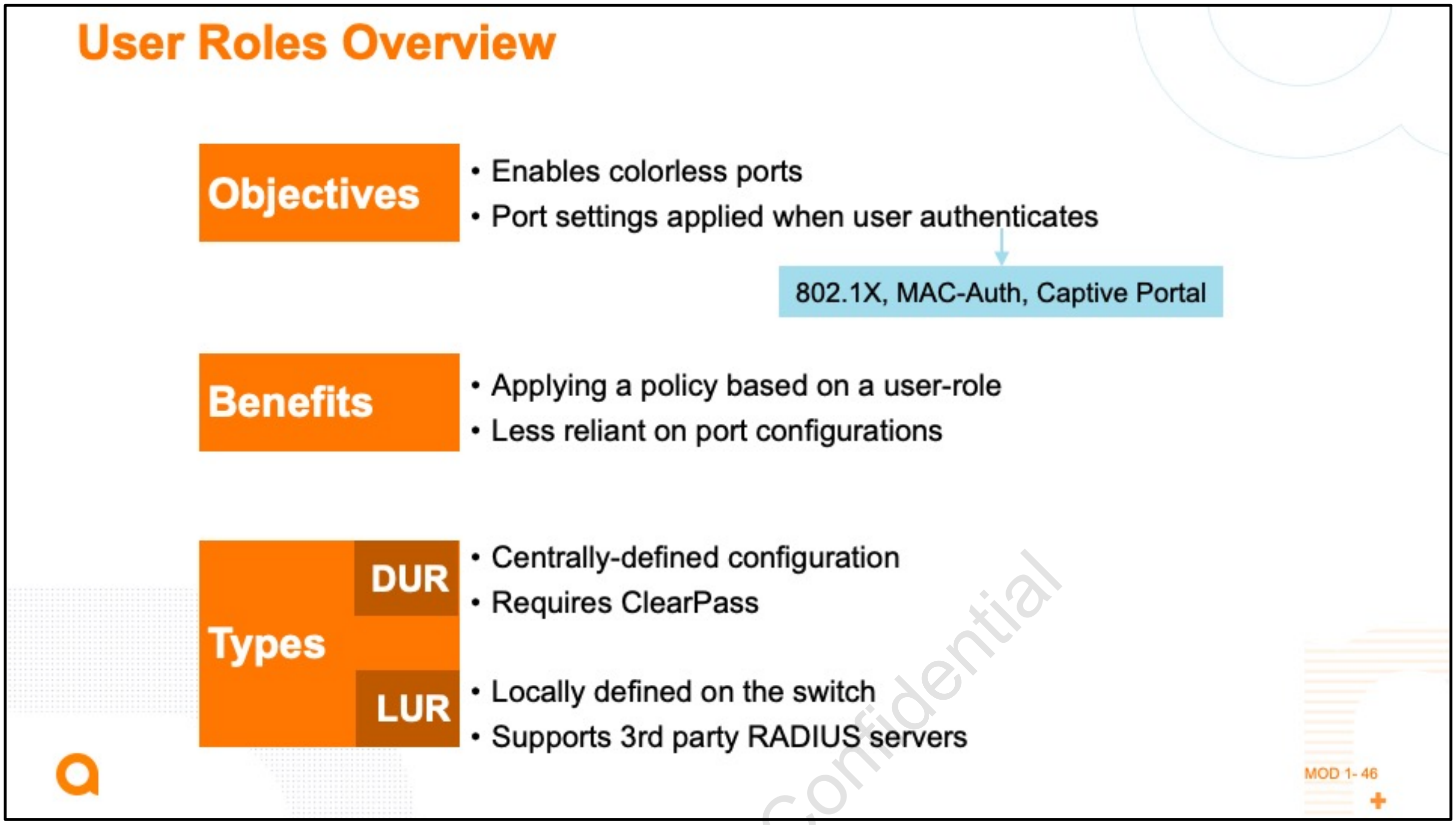
PEAP gives you less control, since you are unaware of the user's device type

If using PEAP, compliment 802.1X authentication with a fingerprinting feature

Even with EAP-TLS, although you can identify the who and the what, you still lack the rick who, what, how, where, when context provided by ClearPass

On a related note, read the note here about how EAP-TLS can be a better choice than PEAP, but use of a fingerprinting feature like this can really improve things.

Important: Using EAP-TLS with certificates gives you much more control over what is connecting to the network - you control which devices get certificates. Using PEAP for authentication gives you less control, since you don't know what device the user connects with. If using PEAP, you should compliment your 802.1X authentication with a fingerprinting feature like that supported by ClearPass. Even with EAP-TLS, although you can identify the who and the what, you still lack the rick who, what, how, where, when context provided by Clearpass.



User Roles Overview

Objectives
- Enables colorless ports
- Port settings applied when user authenticates

802.1X, MAC-Auth, Captive Portal

Benefits
- Applying a policy based on a user-role
- Less reliant on port configurations

Types

DUR
- Centrally-defined configuration
- Requires ClearPass

LUR
- Locally defined on the switch
- Supports 3rd party RADIUS servers

MOD 1- 46

| The benefits of assigning “roles” to users and/or devices is well known within the Aruba wireless world. This has simply been adapted to the wired switch port. Just as with Aruba Wireless, you simplify configuration, grouping policies into a “role” that can be referenced by many device or user types. Additionally, when ClearPass is incorporated, that who, what, how, where, when context can be used to control user access rights.

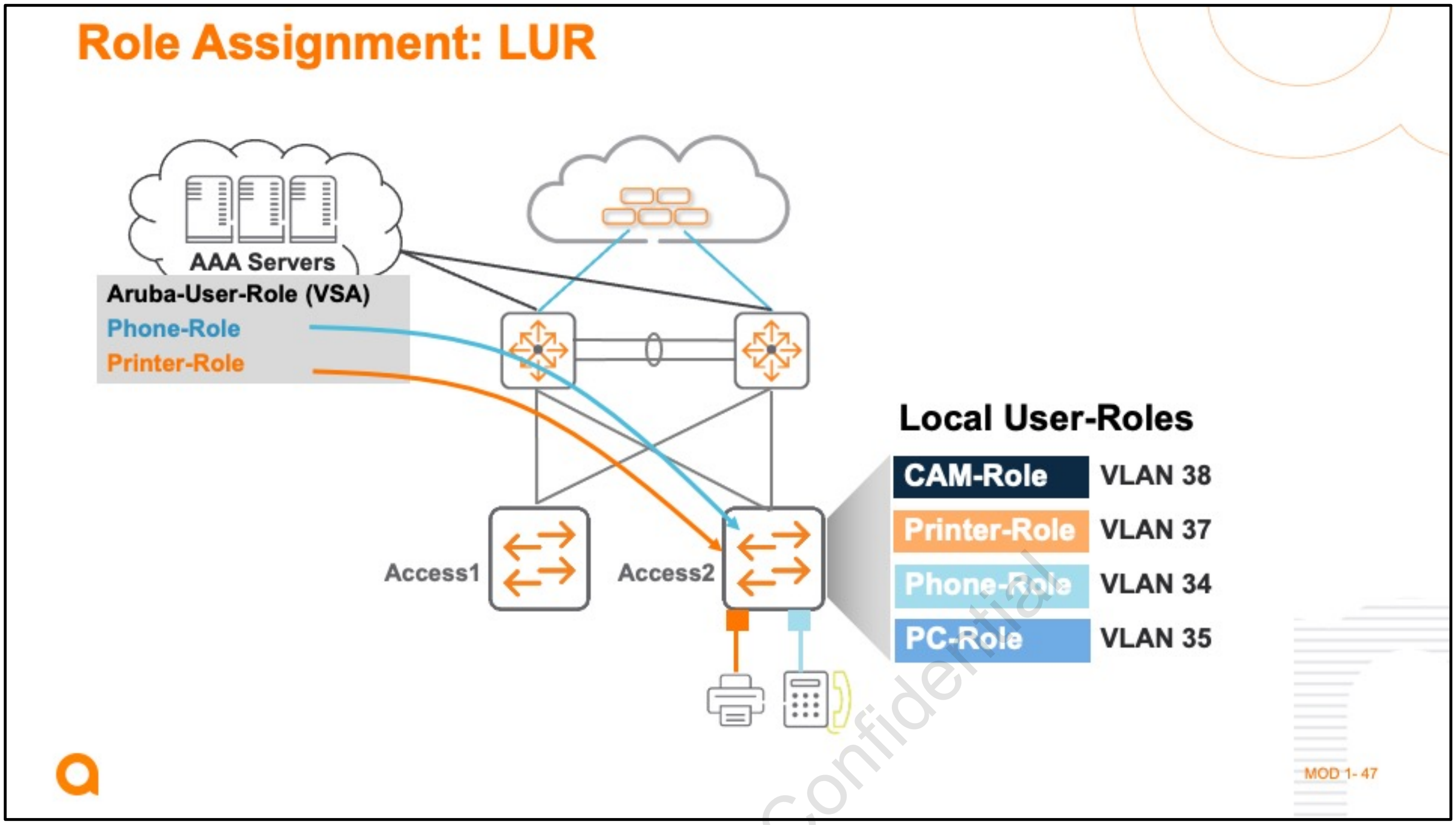
| You can configure roles locally on the switch using a Local User Role (LUR). When a device connects and authenticates successfully, a User-Role name (VSA) is passed to the switch, then the switch will apply that User-Role to the device.

| Or configure Downloadable User Roles (DUR) on a ClearPass server. The ClearPass HPE-CPPM-Role VSA is used with HTTPS to transfer the role to the switch.

A role is a role. At a minimum they dictate the tagged or untagged VLAN assignment, and if the traffic is locally-switched or tunneled back to a Mobility Controller. Optionally a role can also assign a policy (ACL/QOS), reauthentication timers, and a captive portal redirect. The same syntax is used whether pre-defined on a switch (local-roles) or downloaded from ClearPass. It must exist on the switch before it can be applied to a device/user.

Important: the term “role” is used with both the ArubaOS Mobility Controllers (MCs) and the AOS-CX switches. The two should not be confused, since a role on an MC pertains to firewall, bandwidth, and captive portal policies, while a role on an AOS-CX switch pertains to VLAN, QoS, ACL, and other policies. In other words, roles on both devices apply policies, but the policy types differ.

Note: In the current release (10.4), Aruba doesn’t support role download for a gateway (controller) or secondary role to the mobility controller when implementing dynamic segmentation.



Role Assignment: LUR

Local User-Roles

CAM-Role	VLAN 38
Printer-Role	VLAN 37
Phone-Role	VLAN 34
PC-Role	VLAN 35

MOD 1- 47

Remember with LURs, the RADIUS AAA server only passes a VSA that indicates the name of the locally-defined role to be applied.

A user role consists of the following optional parameters:

Ingress user policy: L3 (IPv4 and/or IPv6) ordered list of classes with actions.

captive-portal-profile: Assigns a captive portal profile for this role.

inactivity-timeout: The inactivity timeout period in seconds with a range of 300 to 4294967295 for the authenticated client for an implicit logoff.

reauth-period: Sets the reauthentication period in seconds or 0 to disable.

vlan access: Sets the untagged VLAN ID (native VLAN).

vlan-id-tagged: Sets the tagged VLAN ID.

auth-mode: Sets the configuration in user role to either device-mode or port-mode (client-mode). The following are the attributes:

poe-priority: Specifies the PoE priority for the interface.

mtu: Configures the MTU support for the client.

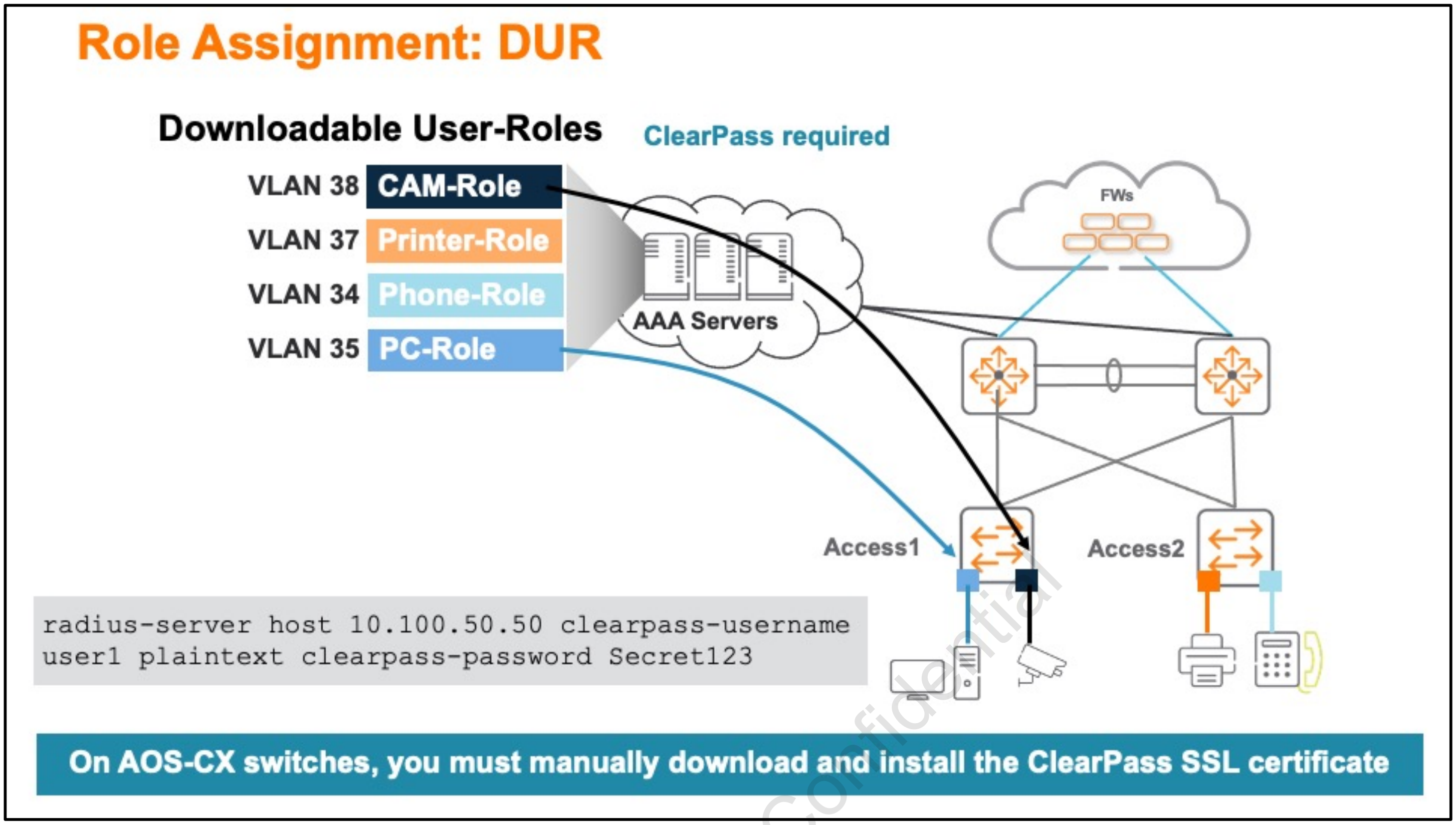
tagged-vlan: Specifies the list of tagged VLANs configured for the interface.

trust-mode: Configures the QoS trust mode for the client.

To create a local user role, use the following configuration:

```
Switch(config)# port-access role <role-name>
Switch(config-pa-role)# description <description>
Switch(config-pa-role)# associate captive-portal-profile <profile-name>
Switch(config-pa-role)# associate policy <policy-name>
Switch(config-pa-role)# vlan access <VLAN-ID>
Switch(config-pa-role)# vlan trunk native <VLAN-ID>
Switch(config-pa-role)# vlan trunk allowed {<VLAN-ID> | all}
Switch(config-pa-role)# trust-mode {dscp | cos | none}
Switch(config-pa-role)# poe-priority {critical | high | low}
Switch(config-pa-role)# auth-mode {client-mode | device-mode}
Switch(config-pa-role)# reauth-period <seconds>
Switch(config-pa-role)# session-timeout <seconds>
Switch(config-pa-role)# client-inactivity time <seconds>
Switch(config-pa-role)# mtu <byte-size>
```

Please note that role names are case-sensitive! Use the show port-access role command to verify your local role's configuration. See the AOS-CX Security Guide for more information to configure the various role properties.



Role Assignment: DUR

Downloadable User-Roles

ClearPass required

```
radius-server host 10.100.50.50 clearpass-username
user1 plaintext clearpass-password Secret123
```

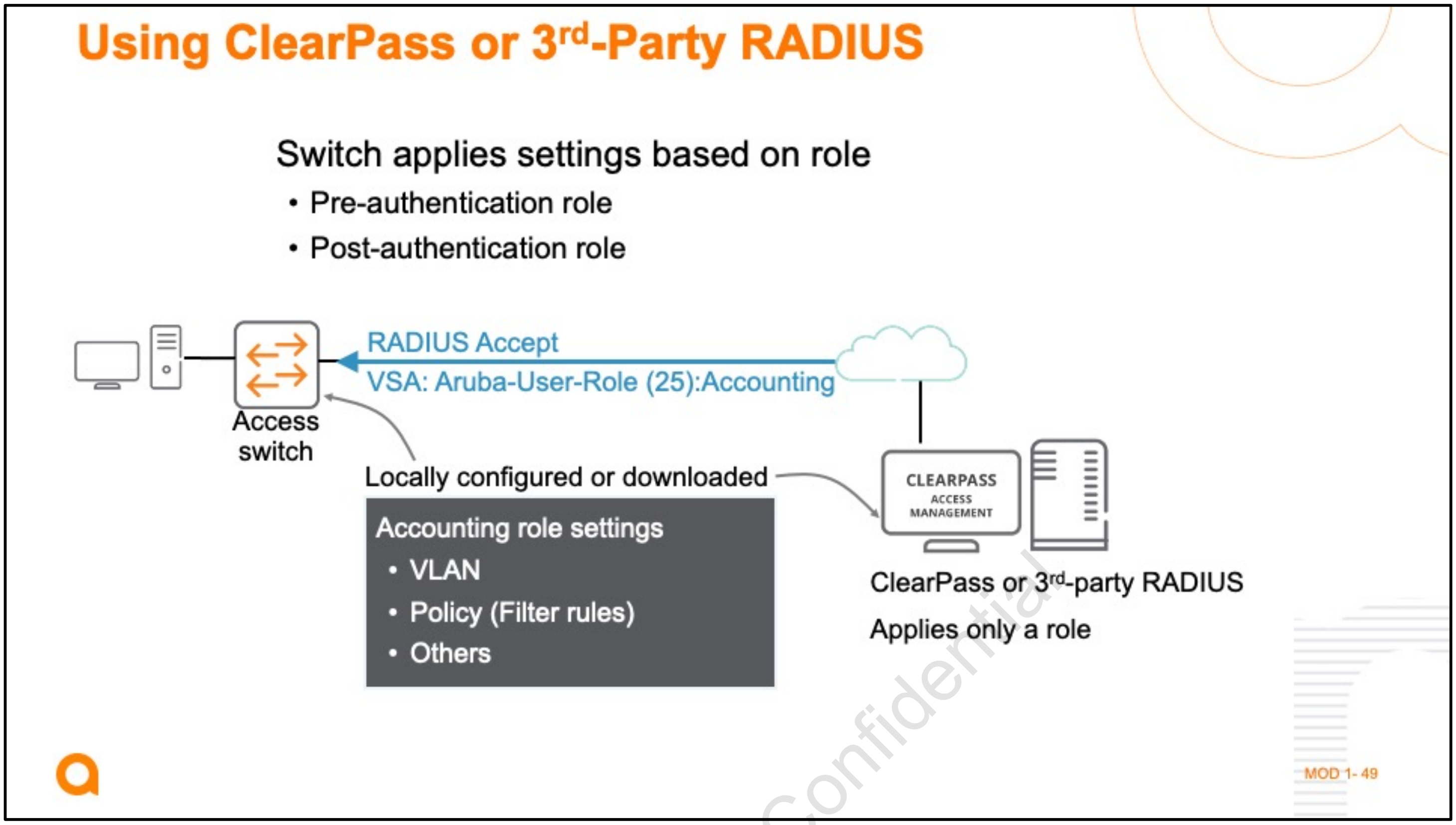
On AOS-CX switches, you must manually download and install the ClearPass SSL certificate

Downloadable user-roles are downloaded from ClearPass Policy Manager to the switch, as opposed to locally configuring roles on switches. The download enables you to setup policies and attributes for a specific user-role which can then be downloaded and stored on the switch.

Configure new users and assign them the same stored version of the user-role in ClearPass, saving the time to reconfigure each user individually.

Use the radius-server host command if the entire user role configuration is to be downloaded from ClearPass Policy Manager server, as shown here.

Note: Because the login process from the AOS-CX switch uses HTTPS, the switch must validate the ClearPass server certificate. On older AOS switches, they would automatically download and install the ClearPass certificate. Unfortunately, in AOS-CX 10.4, this process is not automatic. You must access the ClearPass server, download its SSL certificate, and install it on the AOS-CX switch.



Using ClearPass or 3rd-Party RADIUS

Switch applies settings based on role

- Pre-authentication role
- Post-authentication role

RADIUS Accept

VSA: Aruba-User-Role (25):Accounting

Access switch

Locally configured or downloaded

Accounting role settings

- VLAN
- Policy (Filter rules)
- Others

ClearPass or 3rd-party RADIUS

Applies only a role

MOD 1- 49

You have learned how AOS-CX switches can accept RADIUS VSAs to customize access for authenticated devices.

Some companies prefer to take a different approach. They might use ClearPass to apply user roles to authenticated users based on identities and other criteria. Aruba MCs, acting as RADIUS clients, use these roles to decide how to filter user traffic. Rather than create new profiles with the RADIUS VSAs for dynamic VLANs and ACLs, you might prefer to have ClearPass simply send a role to AOS-CX switches too.

Every client is associated with a role. Roles associate with a set of attributes for authenticated clients (clients with authentication configuration) and unauthenticated clients, applied to each user session. On a switch, roles are always enabled by default.

Examples of user roles are:

Employee—All access

Contractor—Limited access to resources. Each user role determines the client network privileges,

frequency of reauthentication, applicable bandwidth contracts, and other permissions.

Guest—Browse Internet

The VSA that AOS-CX switches accept from RADIUS servers like ClearPass has these properties:

Type: RADIUS vendor-specific for Aruba

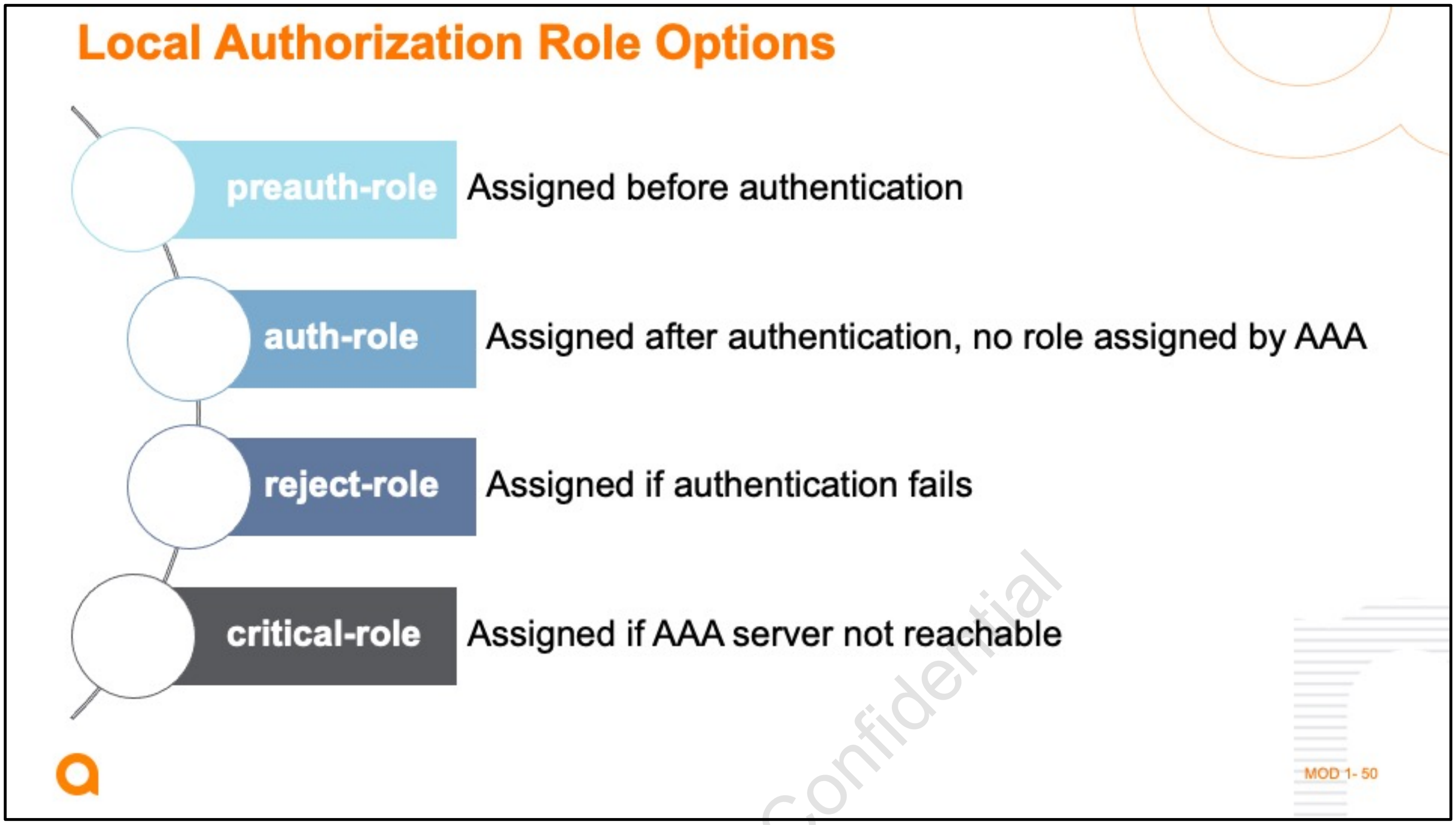
Name: Aruba-User-Role

ID: 25

Value: <role name>

Note: The role name is case-sensitive!

You can define the user role locally. This approach moves the control over customizing settings from RADIUS server administrators to infrastructure administrators, which some companies might prefer. The User Derived Role is applied when the roles are enabled. UDR will have the same precedence order as the authentication type (802.1x, MAC authentication). UDRs are locally configured on the switch using the port-access role command. The configuration of UDRs is beyond the scope of this course. See the Security Guide for more information to configure the port-access role profile locally.



Local Authorization Role Options

preauth-role	Assigned before authentication
auth-role	Assigned after authentication, no role assigned by AAA
reject-role	Assigned if authentication fails
critical-role	Assigned if AAA server not reachable

MOD 1- 50

The figure shows four commonly used role types. Click on each role type to explore.

Pre-auth role

The pre-authentication (pre-auth) role allows a device, such as an IP Phone, to have network access before the device is authenticated. The pre-auth role can also be used to allow DHCP, DNS, and the web redirect for captive portal authentications.

The pre-auth role is triggered when a MAC-based client is connected to an Aruba switch before being authenticated by the RADIUS Server. To provide network connectivity for devices, they must be assigned a VLAN. Two new VLANs are created for pre-auth role functionality, one for voice traffic and one for data traffic. Pre-auth role VLANs can be configured on the switch individually or within a user-role. Devices that can be connected to the switch without authentication are divided into two categories:

Devices that send voice traffic.

Devices that send data traffic.

Note: Either one of pre-auth role VLAN (voice and/or data) or a pre-auth role can be configured for a port. However, both a VLAN and role cannot coexist for an interface. Initial traffic on the port is restricted only by ACLs configured for the port or for VLANs or ACLs in the role.

To define a pre-authentication role, configure the following:

```
Switch(config)# interface <interface-ID>
Switch(config-if)# aaa authentication port-access preauth-role <role-name>
```

Impact of pre-auth role on existing features include the following:

Unauthenticated devices: Configuring pre-auth role VLAN will change the behavior of unauthenticated devices. Normally, authentication-enabled ports will not provide unauthenticated client any network access until the device is authenticated by the RADIUS Server. With pre-auth role VLAN configured, the client will be assigned to the pre-auth role VLAN until the RADIUS server authenticates the device. Unauthenticated clients will be placed into the VLAN specified in the pre-auth role. After authentication by the RADIUS server, the client will be placed into the VLAN specified by the RADIUS authentication command string or as specified in the RADIUS authentication accept string.

LLDP-Bypass: When LLDP-bypass is enabled on the switch, Aruba APs are not authenticated. Therefore, the pre-auth role VLAN is not applicable.

Bypass using device-identity: Pre-auth role VLAN is not applicable to VoIP devices because they do not need authentication. It is applicable to PCs which need authentication.

ACLs applied on an Interface: If an ACL rule is applied on an interface which is part of a pre-auth role VLAN, traffic coming through that interface will be affected. Traffic will be affected based on the rule in the ACL.

ACLs applied on a VLAN: If an ACL rule is applied on a pre-auth role VLAN, traffic entering that VLAN will be affected. Traffic will be affected based on the rule in the ACL.

Rate-limiting on an interface: If the traffic is rate-limited on an interface as part of a pre-auth role VLAN, the traffic will be impacted. The traffic will be affected based on the rule in the rate-limiting configuration command.

Authenticated or rejected clients: Clients which are either authenticated or rejected by the RADIUS server are given different VLANs. These clients are moved from pre-auth role to new VLANs based on authentication by the RADIUS Server.

MAC pinning: Clients whose MAC addresses are pinned and have undergone authentication will always be treated as authenticated. Pre-auth role VLAN is not applicable in this scenario.

Effect of RADIUS tracking on pre-auth role: If RADIUS tracking is enabled and no RADIUS server is available for authentication, the port will be changed from a pre-auth role VLAN to a critical VLAN. The time taken to move from pre-auth role VLAN to critical VLAN depends on the time it takes for RADIUS tracker to inform the subsystem.

Auth-Role

When implementing AAA, a AAA server will authenticate the user. If the AAA server doesn't reply with a role name, or the role name send in the AAA request doesn't exist on the switch (case-sensitive), then a default role can be assigned to the user if the following has been configured.

To define a pre-authentication role, configure the following:

```
Switch(config)# interface <interface-ID>
Switch(config-if)# aaa authentication port-access auth-role <role-name>
```

Reject role

Normally, if a user fails authentication, they will have no access beyond the switch port. Optionally, you can assign a reject (failed authentication) role to the user that would give them minimal network access, perhaps to download software to become compliant in order to connect to the network, like an 802.1X supplicant. To define a reject role, configure the following:

```
Switch(config)# interface <interface-ID>
Switch(config-if)# aaa authentication port-access reject-role <role-name>
```

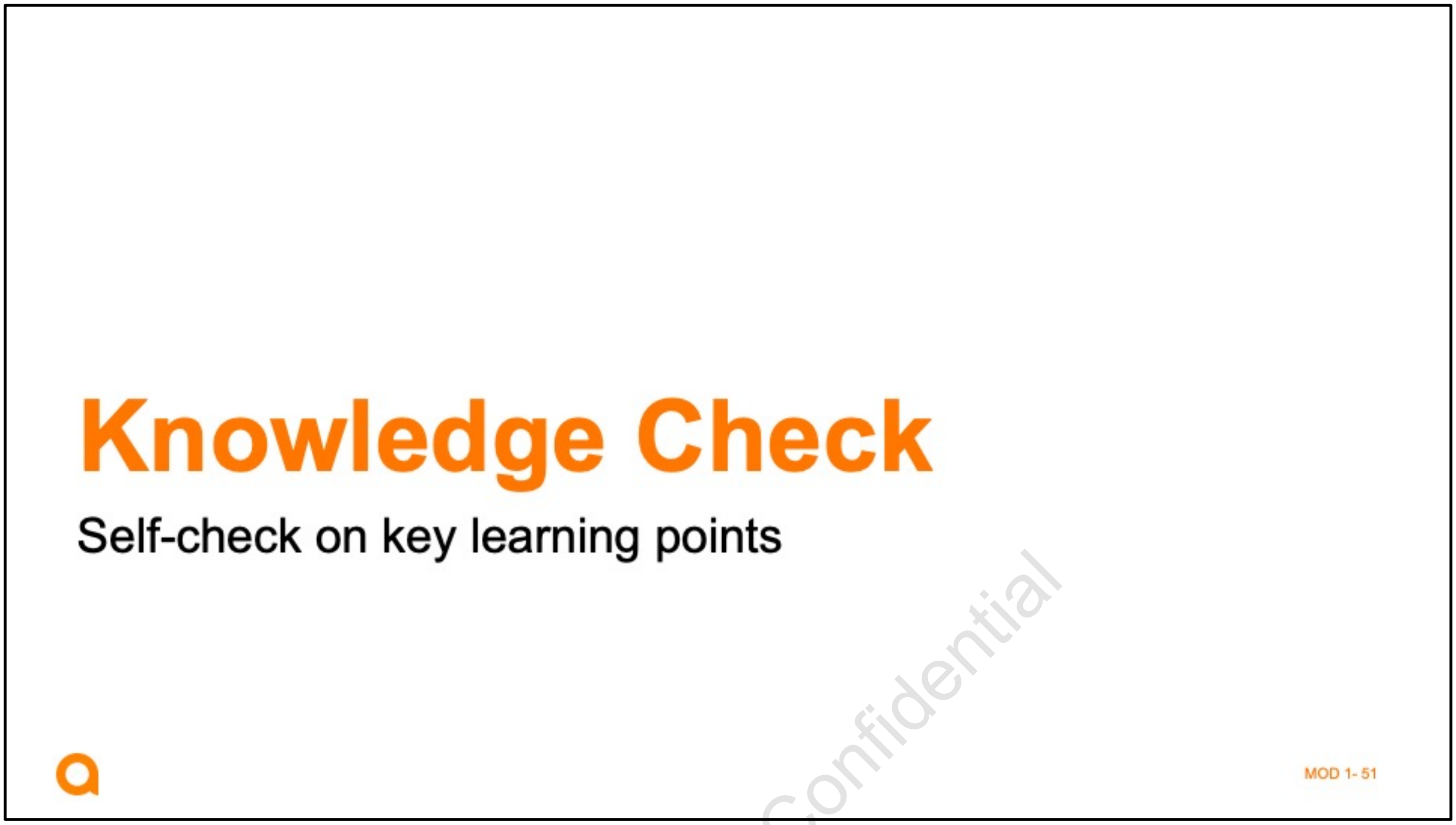
Critical role

A critical role can be assigned to a user when the switch sends an AAA request(s) to an AAA server(s), but no response is received, perhaps because there is a network issue preventing communication between the two sets of devices. By default, the authentication would fail and the user would be denied access. However, you can override this behavior by defining a critical role for the interface, like this:

```
Switch(config)# interface <interface-ID>
Switch(config-if)# aaa authentication port-access critical-role <role-name>
```

If the AAA server(s) is not reachable, then the critical role is applied. This allows important devices, like VoIP phones, for example, to stay connected to the network at all times.

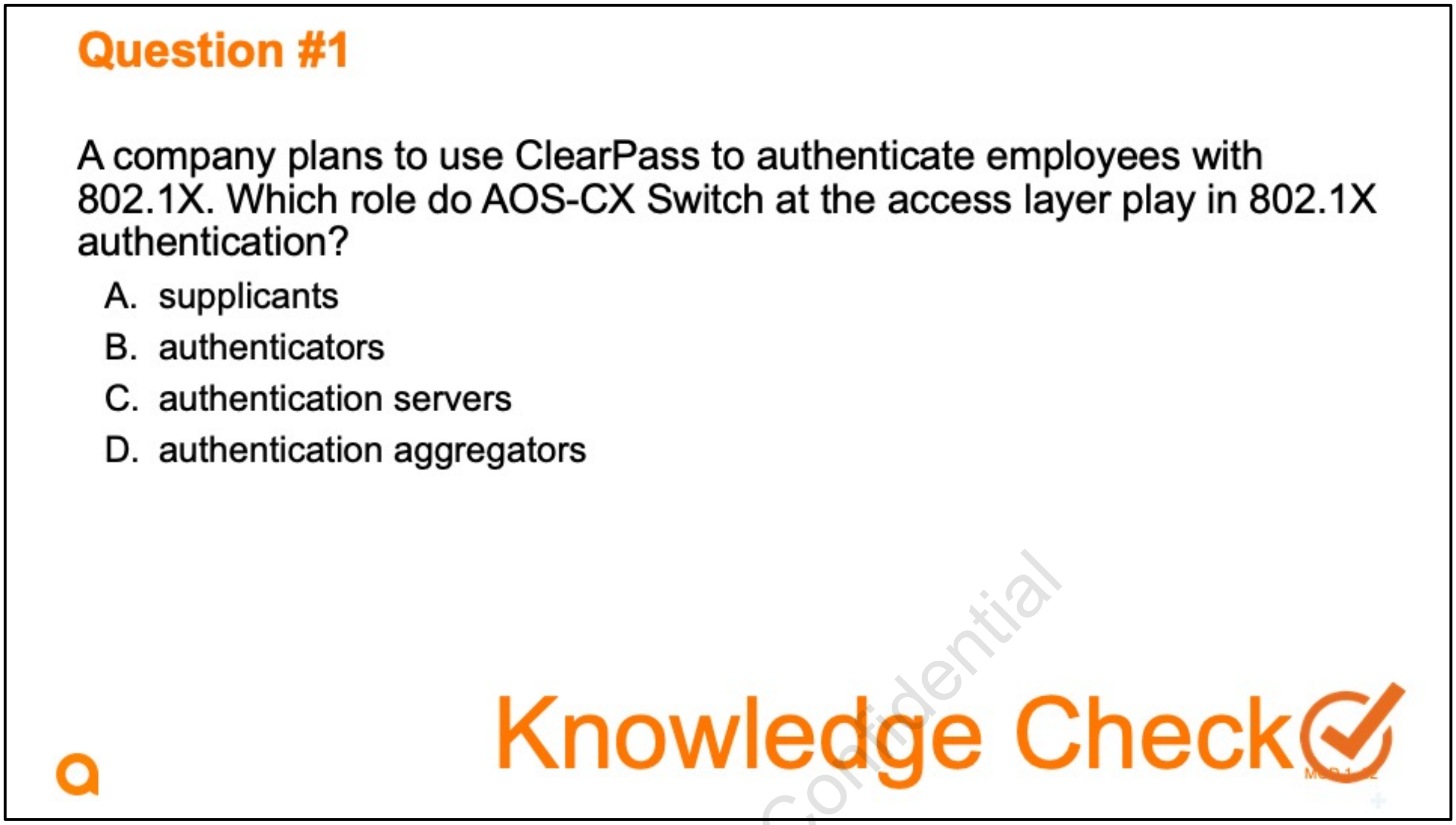
Once an AAA server is reachable, however, the device(s) connected to the interface must be authenticated, at which point the critical role is not used.



Knowledge Check

Self-check on key learning points

MOD 1- 51



Question #1

A company plans to use ClearPass to authenticate employees with 802.1X. Which role do AOS-CX Switch at the access layer play in 802.1X authentication?

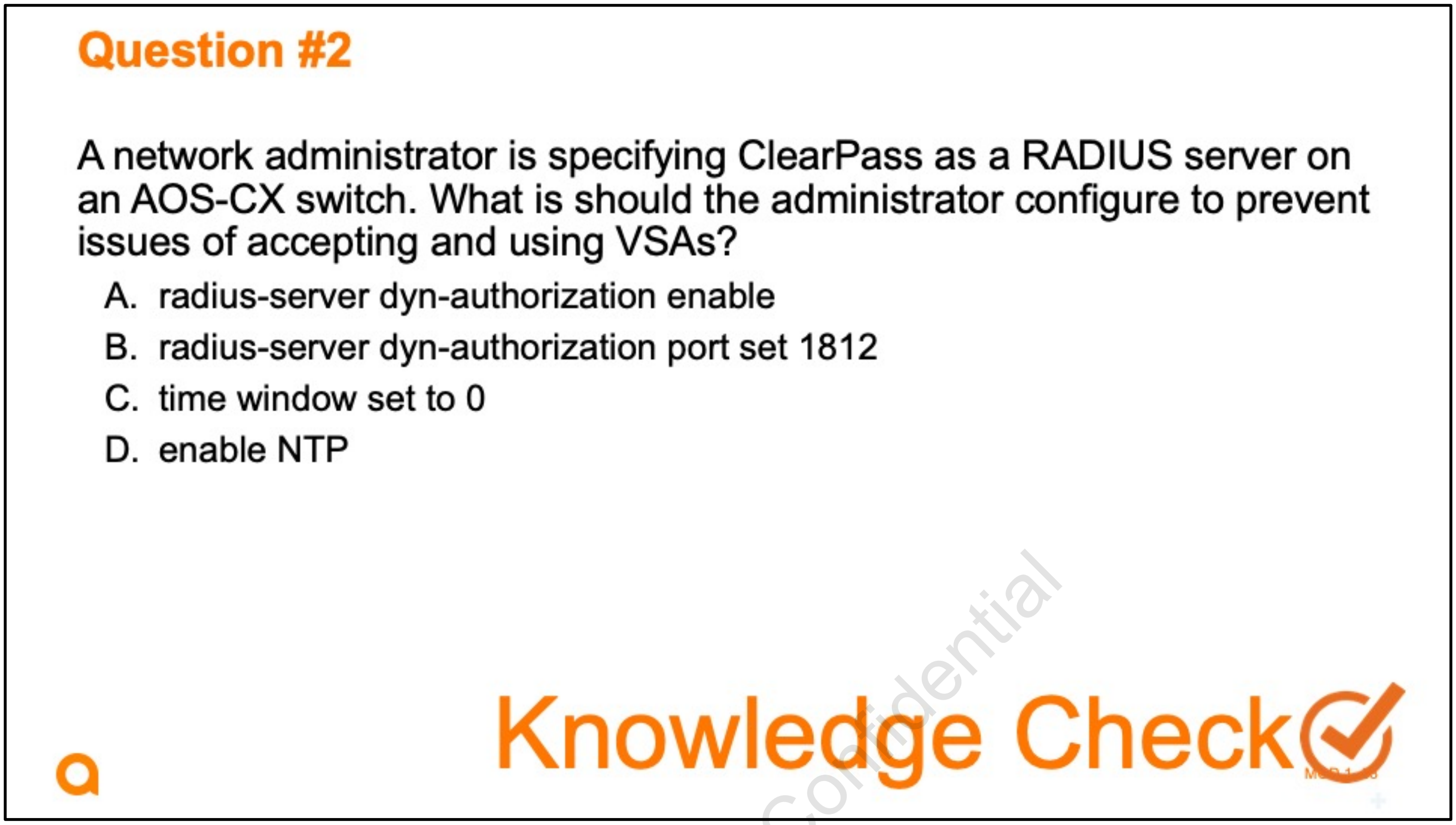
A. supplicants

B. authenticators

C. authentication servers

D. authentication aggregators

Knowledge Check



Question #2

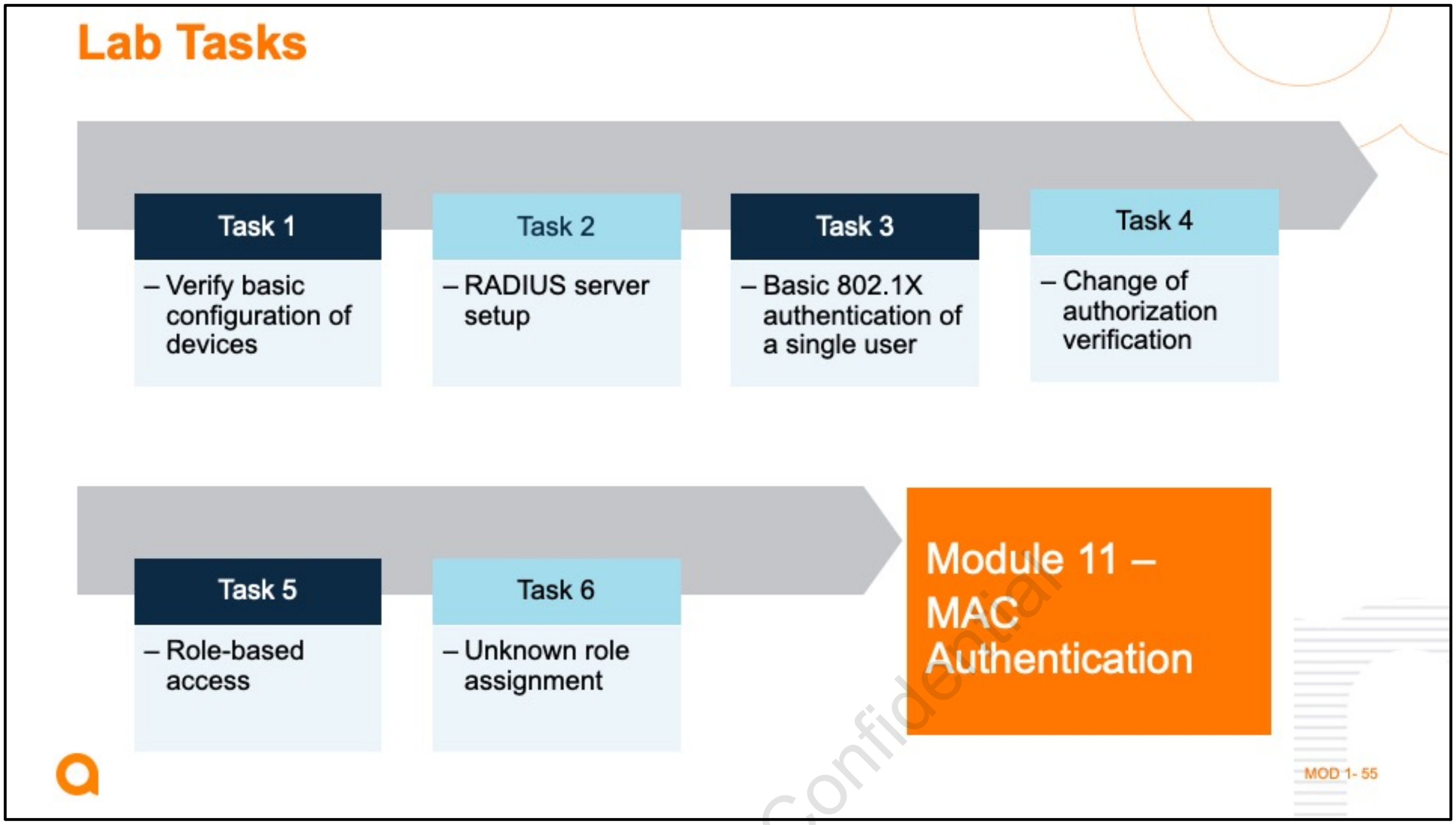
A network administrator is specifying ClearPass as a RADIUS server on an AOS-CX switch. What is should the administrator configure to prevent issues of accepting and using VSAs?

A. radius-server dyn-authorization enable
B. radius-server dyn-authorization port set 1812
C. time window set to 0
D. enable NTP

Knowledge Check

Lab Activity

Lab 6.3



Lab Tasks

Task 1

– Verify basic configuration of devices

Task 2

– RADIUS server setup

Task 3

– Basic 802.1X authentication of a single user

Task 4

– Change of authorization verification

Task 5

– Role-based access

Task 6

– Unknown role assignment

Module 11 – MAC Authentication

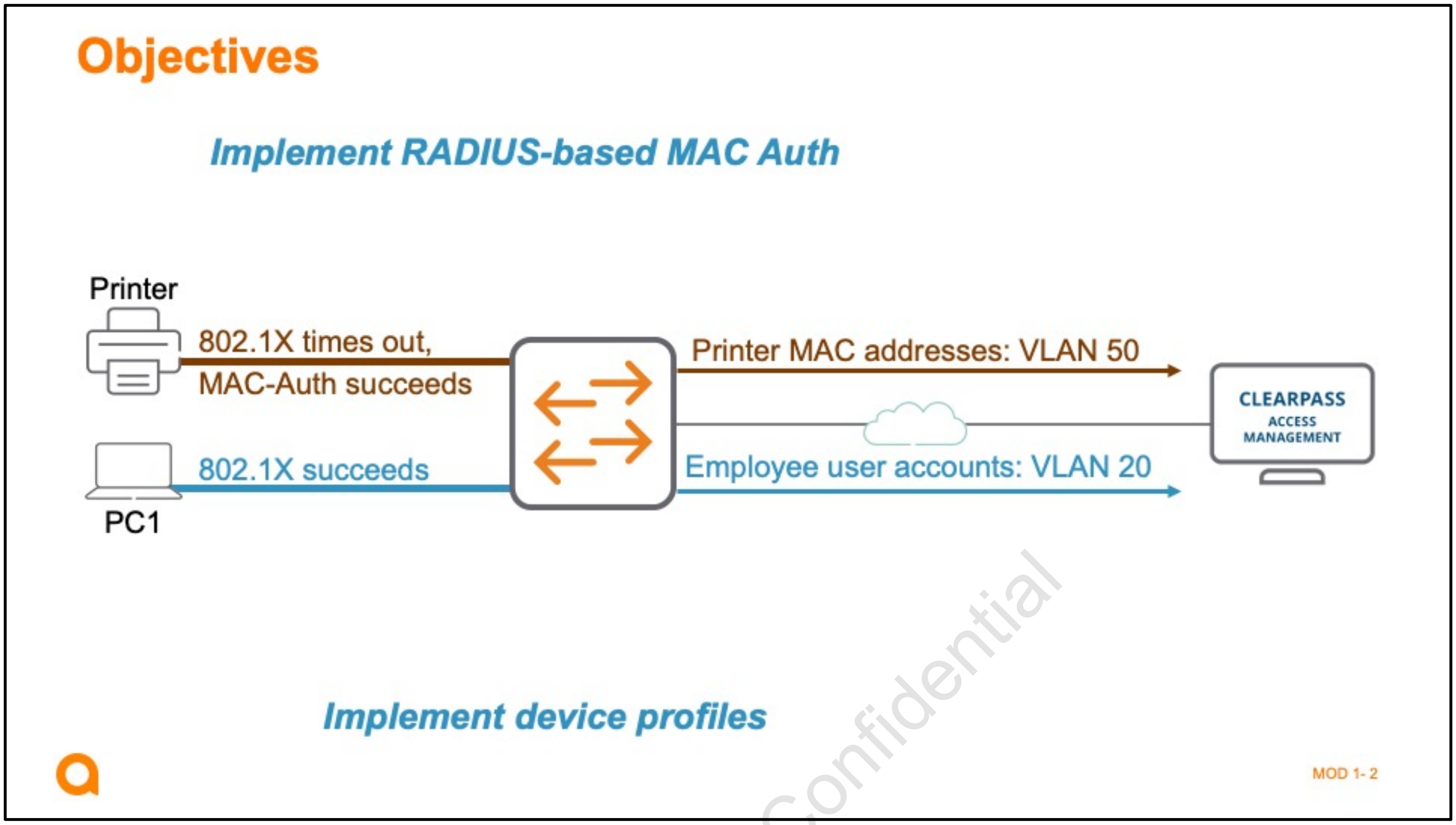
MOD 1- 55

The figure provides a brief review of lab tasks. Please see your lab guide for details.

Rev # 20.21

MAC Authentication

Implementing ArubaOS-CX Switching



Objectives

Implement RADIUS-based MAC Auth

Printer

802.1X times out,
MAC-Auth succeeds

802.1X succeeds

PC1

Printer MAC addresses: VLAN 50

Employee user accounts: VLAN 20

CLEARPASS
ACCESS
MANAGEMENT

Implement device profiles

MOD 1- 2

By the time that you have completed this module, you should be able to:

Implement RADIUS-based MAC Authentication (MAC-Auth) on AOS-CX switch ports

Implement device profiles to provision devices by dynamically assigning the switch port settings based on LLDP or CDP information

Overview

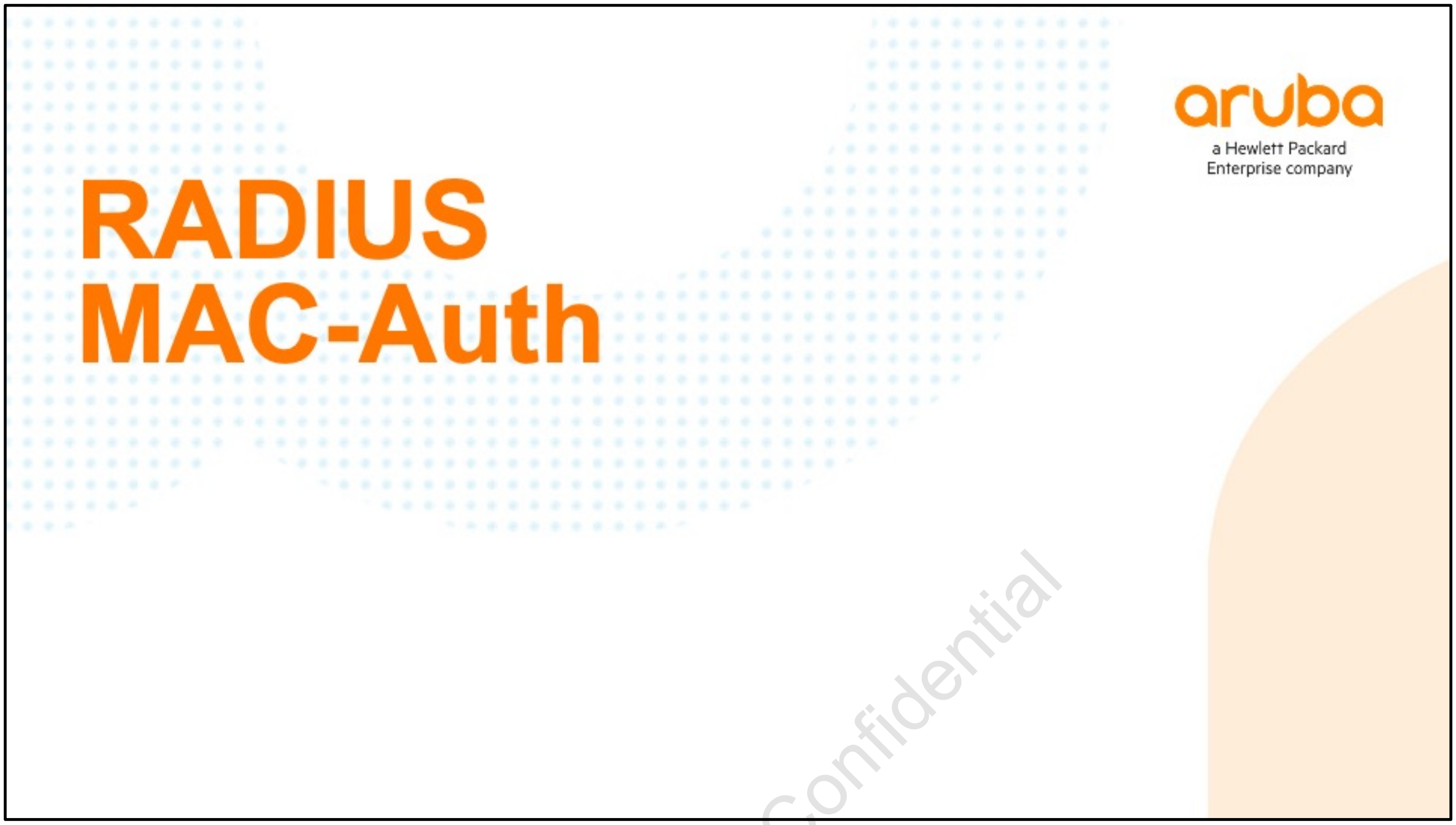
RADIUS MAC-Auth

Authentication Scenarios

Lab Activity

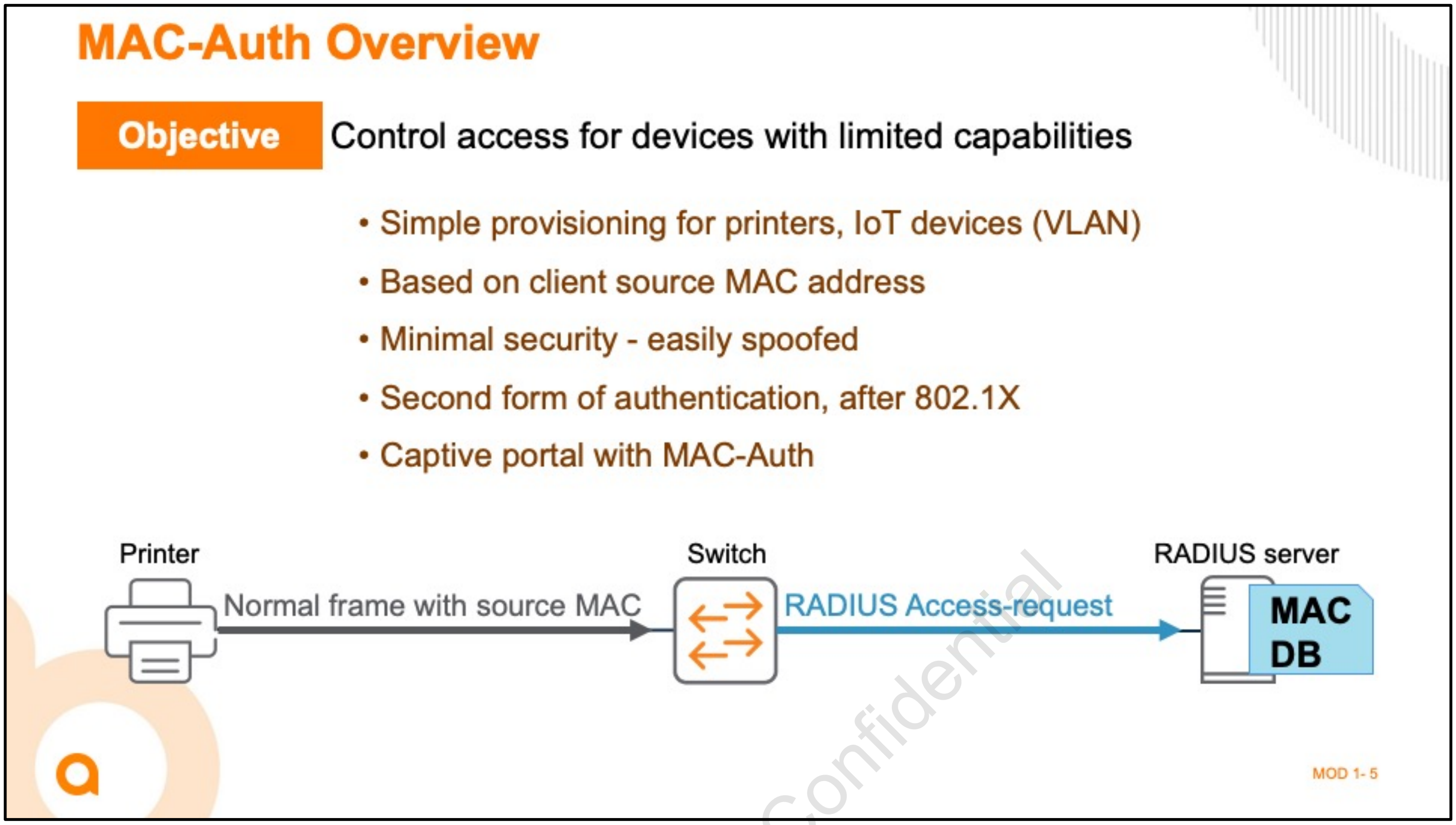
MOD 1- 3

This module teaches you how to implement MAC Authentication (MAC-Auth) technology on AOS-CX switches. Typically MAC authentication is used when devices don't support other methods of authentication, as a second form of authentication, and in combination with captive portal. In this module you'll learn how to implement MAC authentication and how to configure a port where both 802.1X and MAC-Auth are supported on a switch port.



RADIUS MAC-Auth

You will first learn how to implement MAC-Auth with RADIUS-based authentication.



MAC-Auth Overview

Objective Control access for devices with limited capabilities

- Simple provisioning for printers, IoT devices (VLAN)
- Based on client source MAC address
- Minimal security - easily spoofed
- Second form of authentication, after 802.1X
- Captive portal with MAC-Auth

MOD 1- 5

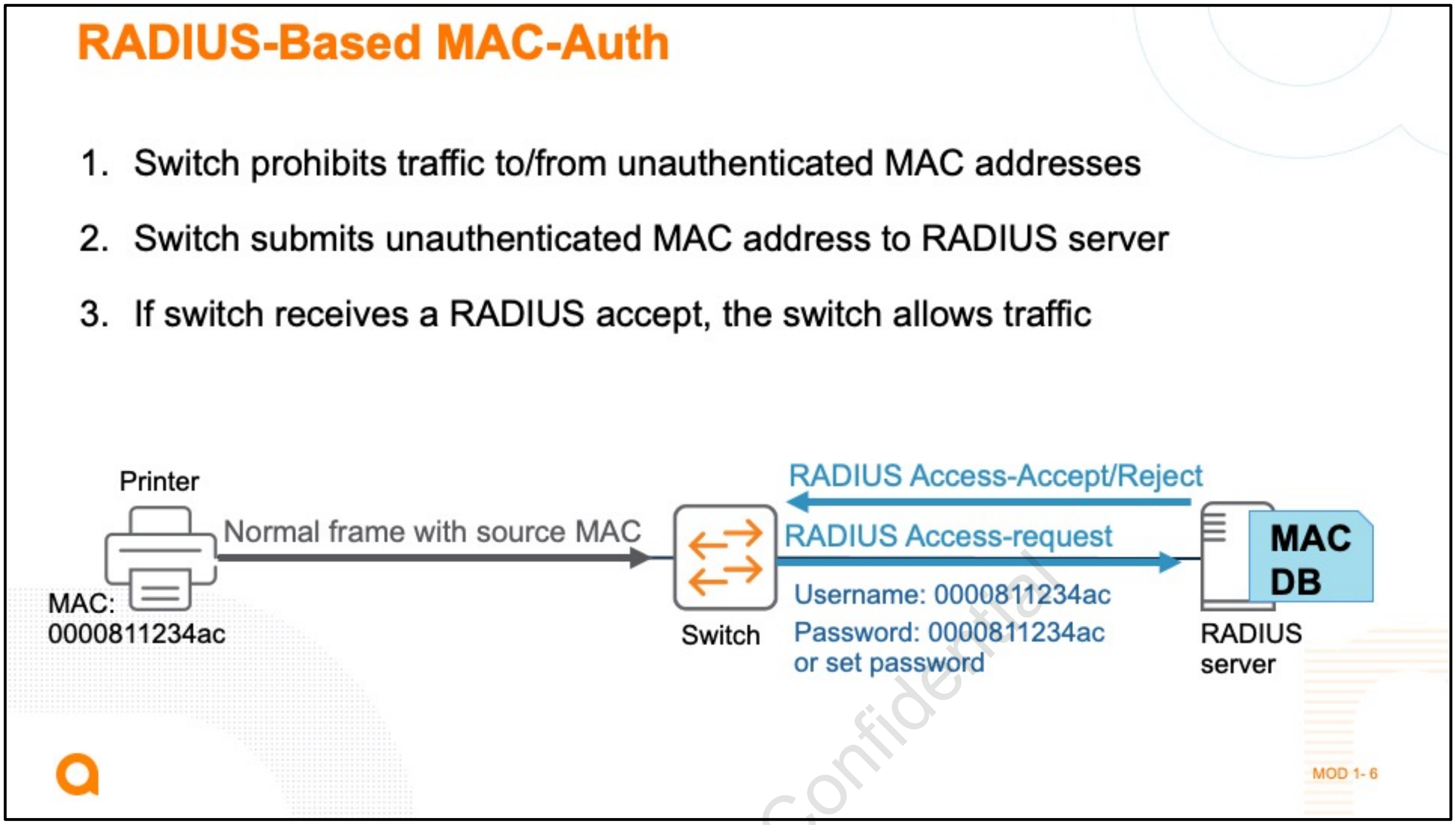
MAC-Auth enables access control for devices with limited capabilities, with simple provisioning for things like printers and IoT devices that may not have an extensive user interface or lack support for 802.1X. In addition to permitting a device network access based on its group, the authentication server can assign the dynamic settings, such as a VLAN ID and QoS priority, to device's session.

These devices simply connect and send some normal Ethernet frame, which of course includes their own MAC address as the source. The switch harvests this MAC address and creates a RADIUS Access-Request message to the RADIUS server. No client configuration required.

When you need rigorous security, you should use 802.1X, whenever possible, since it supports username/password and digital certificate authentication at layer-2. Often it is trivial for a hacker to discover MAC addresses used in your system, whether by eavesdropping or accessing a legitimate device. The hacker can then easily spoof an authorized MAC address. MAC authentication might also be more difficult to manage than other forms of authentication, requiring directory or domain administrators to create accounts for each MAC address on each device, whereas 802.1X can use the user accounts already in the directory or domain.

MAC authentication relies on a RADIUS server. No need to configure MAC-addresses on each switch – a single RADIUS server has a MAC DB, for centralized control. You can use up to three RADIUS servers for redundancy, in case access to the primary server fails. It also means the same credentials can be used for authentication, regardless of which switch or switch port is the current access point into the LAN.

The switch acts as a port-access authenticator using a RADIUS server, and the CHAP and PAP protocols. Inbound traffic is processed by the switch alone, until authentication occurs. Some traffic from the switch to an unauthorized client is supported (for example, broadcast or unknown destination packets) before authentication occurs.



RADIUS-Based MAC-Auth

1. Switch prohibits traffic to/from unauthenticated MAC addresses
2. Switch submits unauthenticated MAC address to RADIUS server
3. If switch receives a RADIUS accept, the switch allows traffic

You can enable RADIUS-based MAC-Auth on individual switch ports, to enable the process described here:

The client is initially unauthenticated. The switch does not transmit traffic on the line. Authentication is not triggered until the client begins to send traffic on its own. This traffic is initially dropped while the switch completes the next steps.

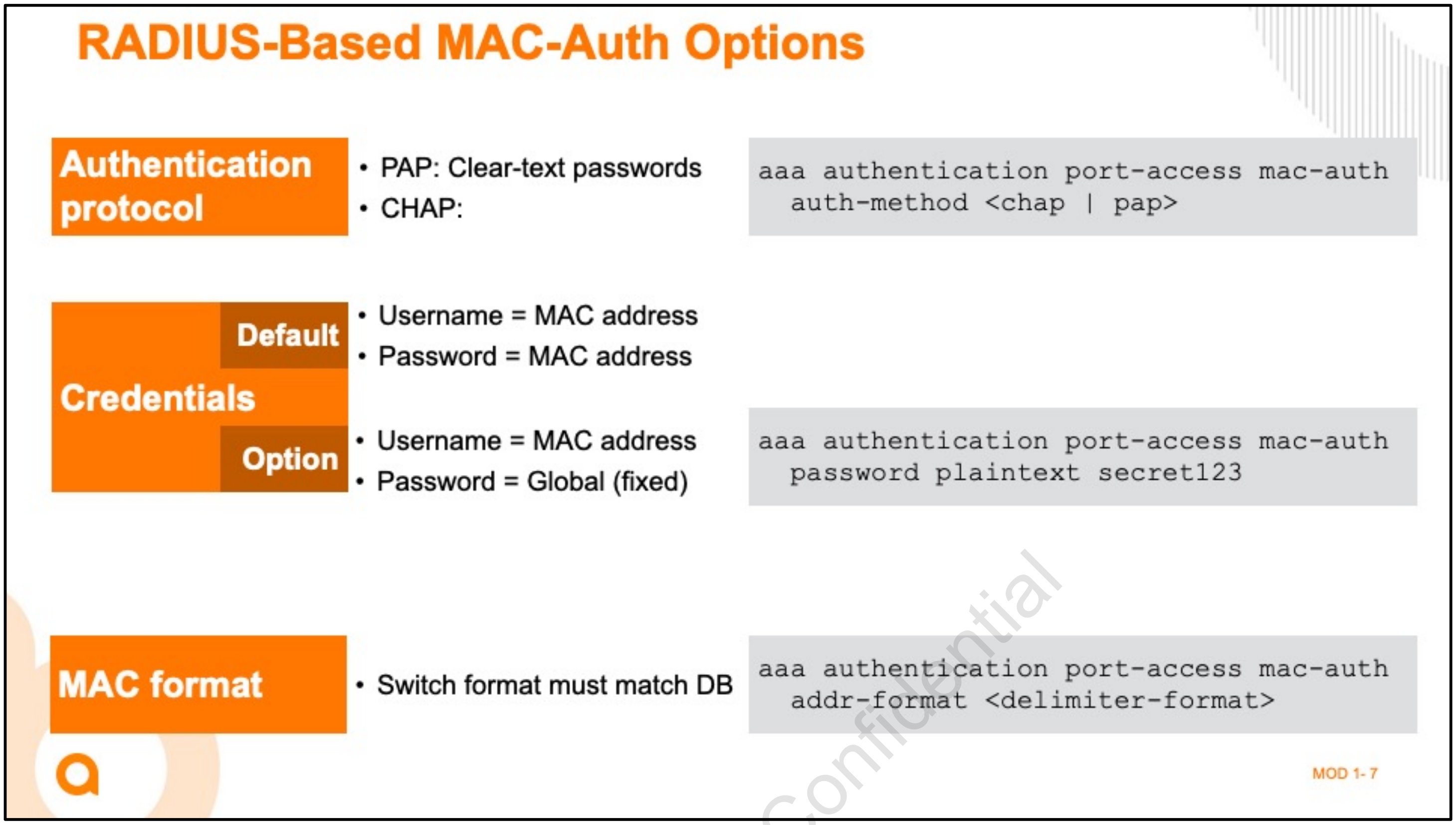
The switch submits credentials to a RADIUS server on the client's behalf. Often these credentials consist of the client's MAC address as the username and password. However, the password can be different from the MAC address.

The RADIUS server authenticates the MAC address against a user account. Only the switch and the server are involved in the exchange of authentication messages, unlike with other port-based authentication methods. The RADIUS server uses its user database—which should have an account for the client's MAC address—and policies to determine whether to grant access. It communicates its decision to the switch in a RADIUS Accept-Access or Reject-Access message.

The switch handles the client according to the RADIUS server's decision.

- If the server accepted access, the switch allows all traffic to and from the client with that MAC address.
- If the server rejected access, the client remains in the unauthenticated state, and the switch handles its traffic as indicated in step 1.

The switch continues to examine incoming frames. If a new client connects, the switch will authenticate that client. The switch also attempts to authenticate the client for which authentication failed again. By default, the switch suppresses authentication requests for such clients over the quiet period—by default, 60 seconds. It then sends another authentication request.



RADIUS-Based MAC-Auth Options

Authentication protocol

- PAP: Clear-text passwords
- CHAP:

```
aaa authentication port-access mac-auth
  auth-method <chap | pap>
```

Credentials

Default

- Username = MAC address
- Password = MAC address

Option

- Username = MAC address
- Password = Global (fixed)

```
aaa authentication port-access mac-auth
  password plaintext secret123
```

MAC format

- Switch format must match DB

```
aaa authentication port-access mac-auth
  addr-format <delimiter-format>
```

Authentication protocol

You have two choices for the protocol that the switch uses to send the credentials to the RADIUS server:

Password Authentication Protocol (PAP)

Challenge Handshake Authentication Protocol (CHAP)

You learned about how RADIUS works with EAP in 802.1X. AOS-CX switches can use PAP or CHAP for MAC-Auth. In this case, the switch acts as the authentication client rather than the client itself, generating the PAP request messages. Note that PAP sends passwords in clear-text, which can pose a security threat.

RADIUS has specific messages for CHAP authentication and most RADIUS servers support this option. You should be aware that the mechanisms of CHAP require passwords to be stored with reversible encryption, which poses a potential security hazard and is not typical for most directories and domains. This limitation would mean, for example, that the directory, domain, or RADIUS administrator who creates the accounts for the clients would have to take special measures to store the passwords with reversible encryption. (For a Microsoft Windows domain, you can edit the group policy object (GPO) and enable reversal encryption in the computer policies for account settings.)

Configure PAP or CHAP authentication as shown in the figure.

Credentials

You must decide the precise credentials that switches will send so that the RADIUS, directory, or domain administrator can set up the accounts accordingly.

By default, AOS-CX switches send the MAC address for both the username and password. You can set up a global password, which the switch uses for the password instead of the device’s MAC address. The switch itself places the password in the authentication request. No configuration is required on the client. The switch continues to use the client’s MAC address for the username, so every client requires a user account on the RADIUS server (or background directory).

The password option is intended to add a bit of security to the solution. However, keep in mind. Hackers can still spoof an authorized MAC address, connect to the switch, and gain access to the network because the switch adds the password on its end. Remember: MAC-Auth provides basic security; you should use 802.1X when you require more rigorous security. Configure a global password for all MAC addresses as shown in the figure.

MAC address format

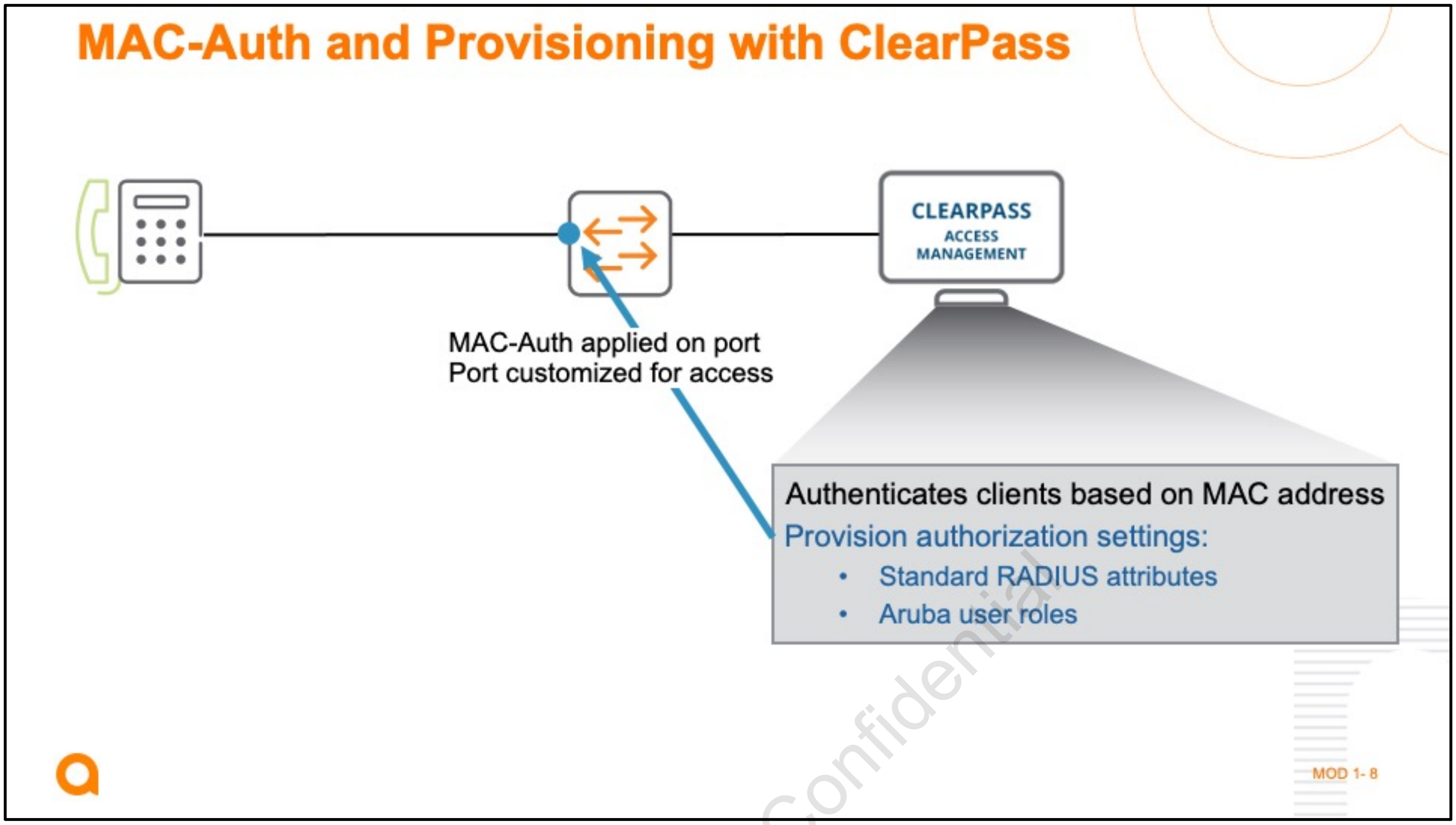
When you are using MAC addresses for the username, password, or both, you must ensure that the switch sends the credentials with the same format used in the user accounts. Configure the delimiter format as shown in the figure. The default is no-delimiter. The table below shows the options and the syntax for configuring a specific format for AOS-CX switches.

Note: Accounts in Microsoft Active Directory (AD) domains cannot use colons, so choose a format that does not include colons if you use AD.

Table 11-1: MAC address formats

Format	AOS-CX switch parameter
aabbccddeeff	no-delimiter
aabbcc-ddeeff	single-dash
aa-bb-cc-dd-ee-ff	multi-dash
aa:bb:cc:dd:ee:ff	multi-colon

AABBCCDDEEFF	no-delimiter-uppercase
AABBCC-DDEEFF	single-dash-uppercase
AA-BB-CC-DD-EE-FF	multi-dash-uppercase
AA:BB:CC:DD:EE:FF	multi-colon-uppercase



MAC-Auth and Provisioning with ClearPass

MAC-Auth applied on port
Port customized for access

Authenticates clients based on MAC address

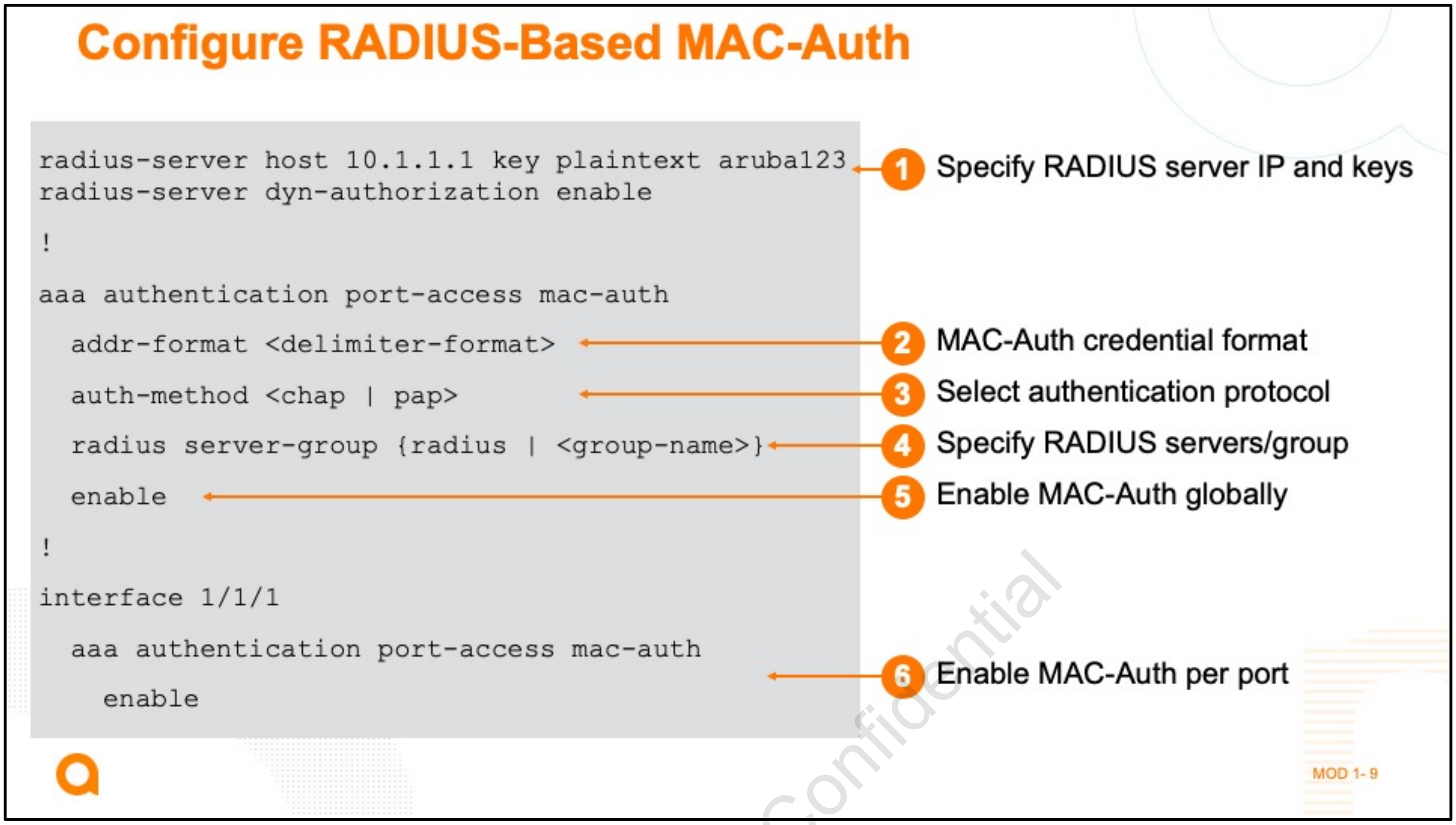
Provision authorization settings:

- Standard RADIUS attributes
- Aruba user roles

Aruba ClearPass can fulfill the role of RADIUS server for MAC-Auth as well as for 802.1X.

ClearPass provides MAC-Auth specific service templates that make it easy to set up. The service can specify profiles to apply to clients based on their MAC addresses and other capabilities. In this way, ClearPass can communicate the correct settings for the connected device, which the switch applies to the port in much the same way as it applies settings with 802.1X. For example, if you are using MAC-Auth to authenticate IP phones, ClearPass policies could assign the correct voice VLAN and QoS priority for the phones’ traffic.

You can provision authorization settings using standard RADIUS attributes, like a VLAN, ACL, or QoS settings, or using Aruba user roles.



Configure RADIUS-Based MAC-Auth

```
radius-server host 10.1.1.1 key plaintext aruba123
radius-server dyn-authorization enable
!
aaa authentication port-access mac-auth
  addr-format <delimiter-format>
  auth-method <chap | pap>
  radius server-group {radius | <group-name>}
  enable
!
interface 1/1/1
  aaa authentication port-access mac-auth
    enable
```

1 Specify RADIUS server IP and keys

2 MAC-Auth credential format

3 Select authentication protocol

4 Specify RADIUS servers/group

5 Enable MAC-Auth globally

6 Enable MAC-Auth per port

Here are the steps to implement MAC-Auth:

Specify your RADIUS server. In this example, you are using only the global group. These commands specify the server and set the proper options for ClearPass. Remember to set the key to match the secret configured for this switch on ClearPass.

Configure the format for the MAC-Auth credentials so that they match the credentials configured on the RADIUS server (or its backend directory). ClearPass often uses the no-delimiter option, which is also the default option on AOS-CX switches.

Choose between PEAP and CHAP for the authentication protocol. Remember that CHAP requires passwords to be stored with reversible encryption, which is not the default for many directories.

Specify the RADIUS server group.

Enable MAC-Auth globally.

Enable MAC-Auth on the individual ports that you want to enforce this type of authentication.

Verification

You can use the following commands to verify the configuration and operation of MAC-Auth:

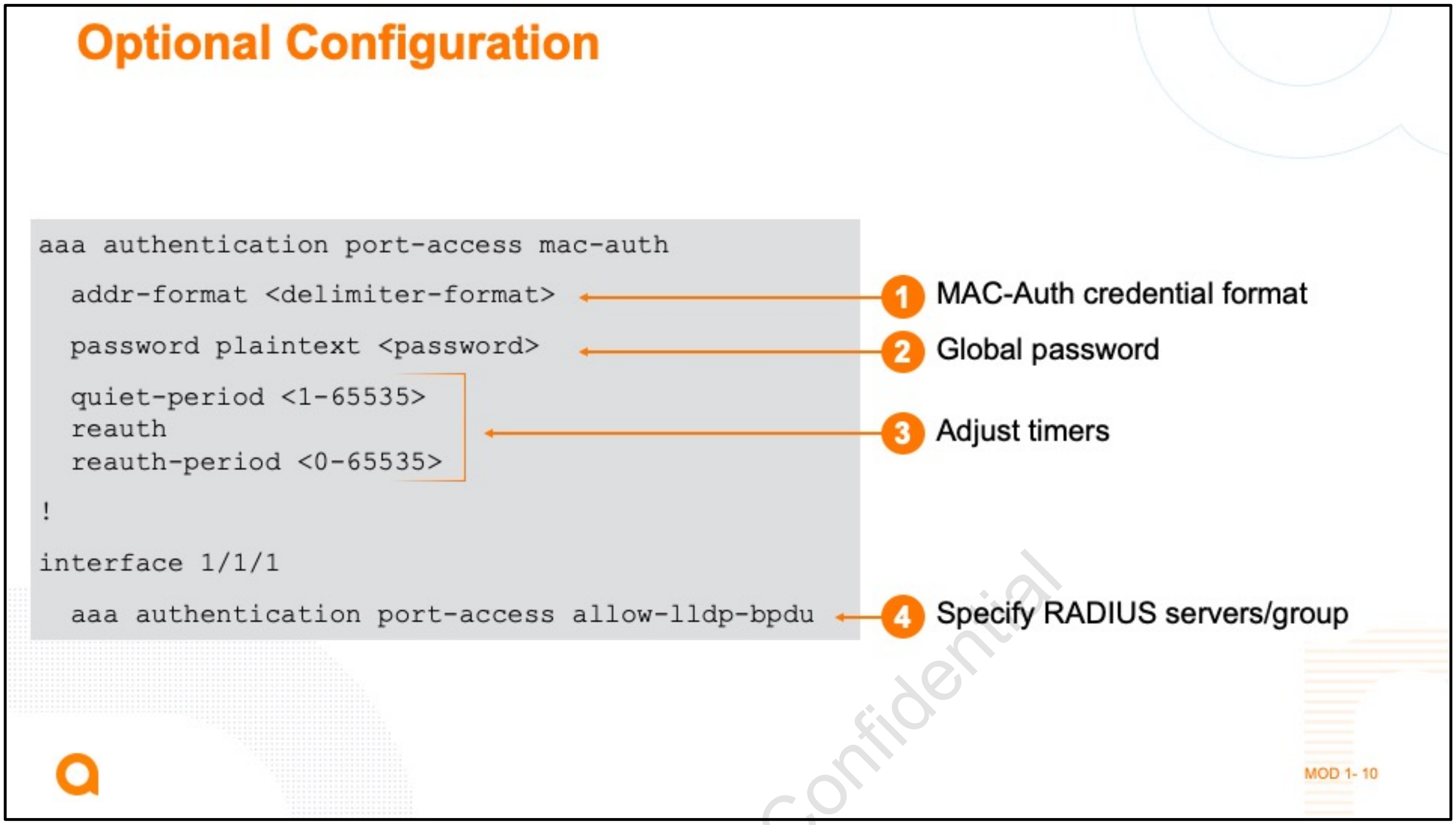
show aaa authentication port-access mac-auth interface [interface name] | all port-statistics: Display information related to MAC authentication ports on the switch.

show aaa authentication port-access mac-auth interface all client-status: This command display MAC authentication clients status on the interface.

clear mac-auth statistics [interface]: Clear the MAC authentication statistics associated with the port and all the MAC authentication state machines attached to this port.

show tech mac-auth [local-file]: Displays output of the following commands:

- show aaa authentication port-access mac-auth interface all port-statistics
- show aaa authentication port-access mac-auth interface all client-status
- show aaa authentication port-access interface all client-status
- diag-dump mac-auth basic



Optional Configuration

```
aaa authentication port-access mac-auth
  addr-format <delimiter-format>
  password plaintext <password>
  quiet-period <1-65535>
  reauth
  reauth-period <0-65535>
!
interface 1/1/1
  aaa authentication port-access allow-lldp-bpdu
```

1 MAC-Auth credential format

2 Global password

3 Adjust timers

4 Specify RADIUS servers/group

MOD 1- 10

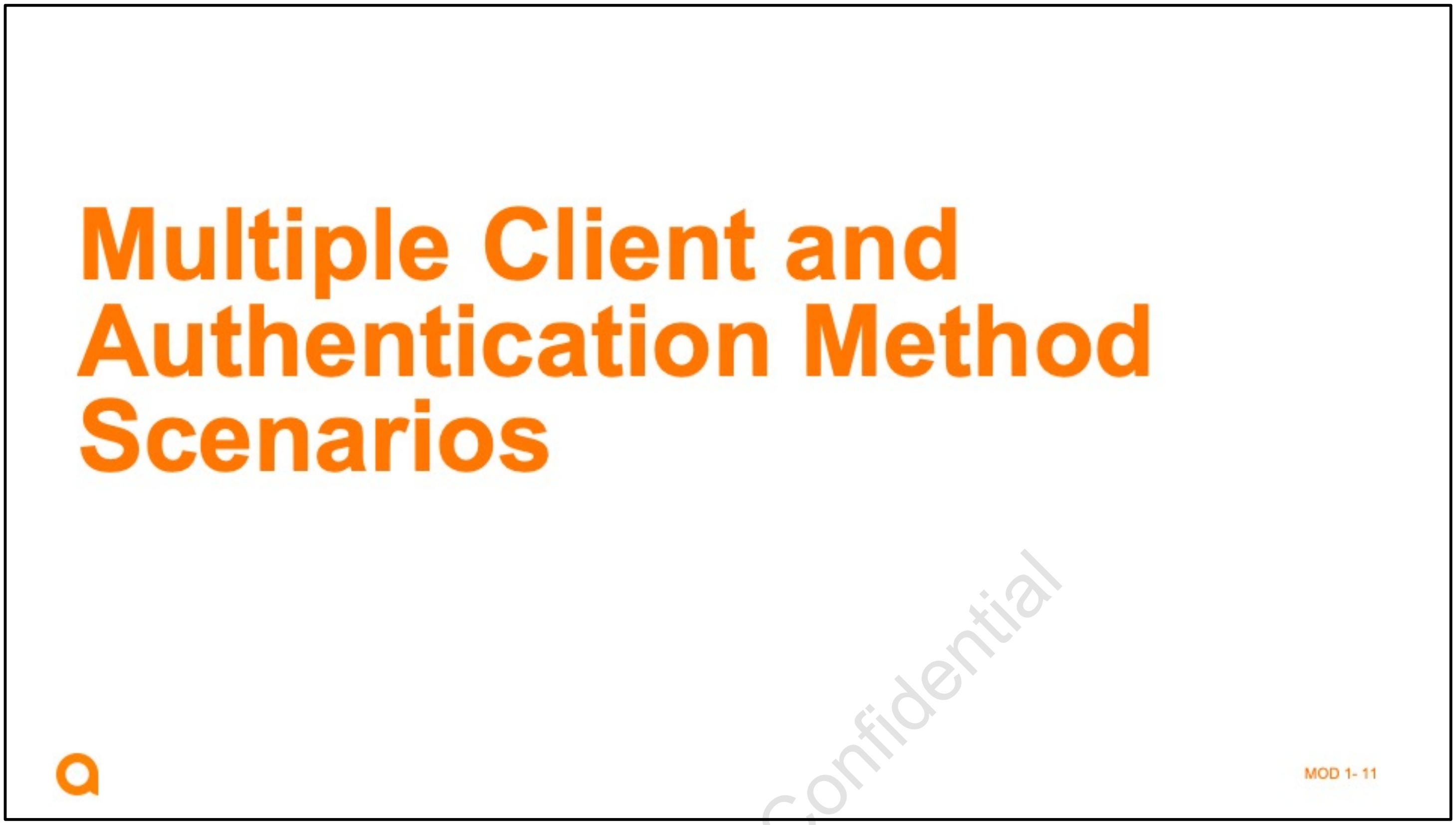
Here are some optional steps for MAC-Auth:

Configure the format for the MAC-Auth credentials so that they match the credentials configured on the RADIUS server or its backend directory. ClearPass often uses the no-delimiter option, which is also the default option on AOS-CX switches.

If you want the switch to send a password, rather than MAC address in the password field, specify the global (fixed) password.

Optionally, you can adjust other settings such as the timers for the quiet period and re-authenticating clients. (If you set the reauth period to 0, the switch does not enforce reauthentication.) All of these timers are set in seconds. You can refer to your switch's Access Security Guide to learn about more MAC-Auth related options.

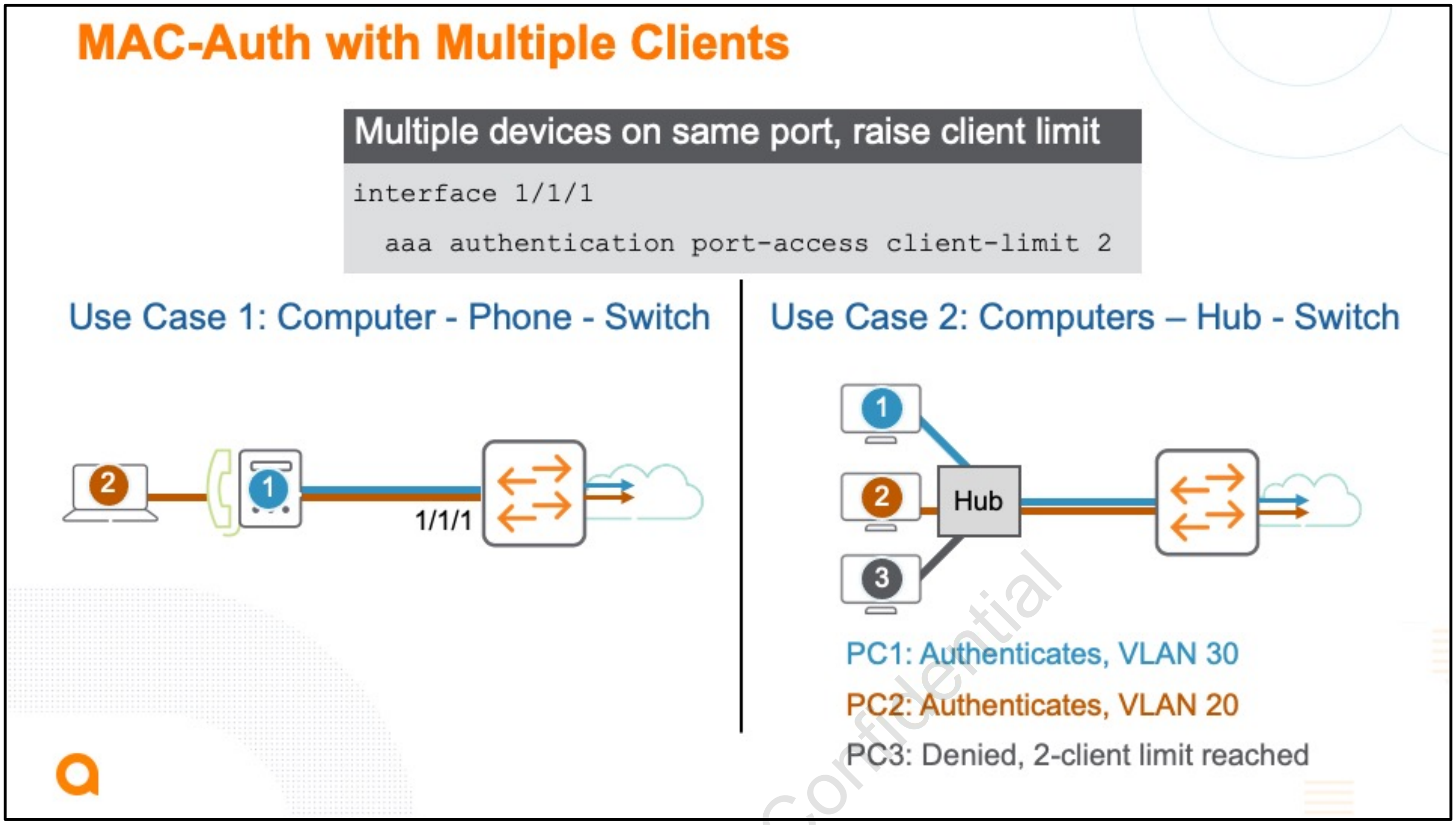
Optionally allow LLDP on the port.



Multiple Client and Authentication Method Scenarios

MOD 1- 11

Next you will look at some more complex scenarios. You will also explore authentication in scenarios in which multiple devices can connect to a single switch. You will learn how to enforce MAC-Auth and 802.1X on the same port.



MAC-Auth with Multiple Clients

Multiple devices on same port, raise client limit

```
interface 1/1/1
  aaa authentication port-access client-limit 2
```

Use Case 1: Computer - Phone - Switch

Use Case 2: Computers – Hub - Switch

PC1: Authenticates, VLAN 30

PC2: Authenticates, VLAN 20

PC3: Denied, 2-client limit reached

Previous scenarios have featured edge ports that support one endpoint at a time, which is very common. However, you might have a Voice over IP (VoIP) deployment, where the VoIP phone connects directly to the switch, and a PC connects to the phone. This is smart, because each user only uses one switch port for both devices. You might apply Local MAC Authentication (LMA) or RADIUS MAC-Auth to the port for IP phones, and 802.1X authentication for the PC, on the same port.

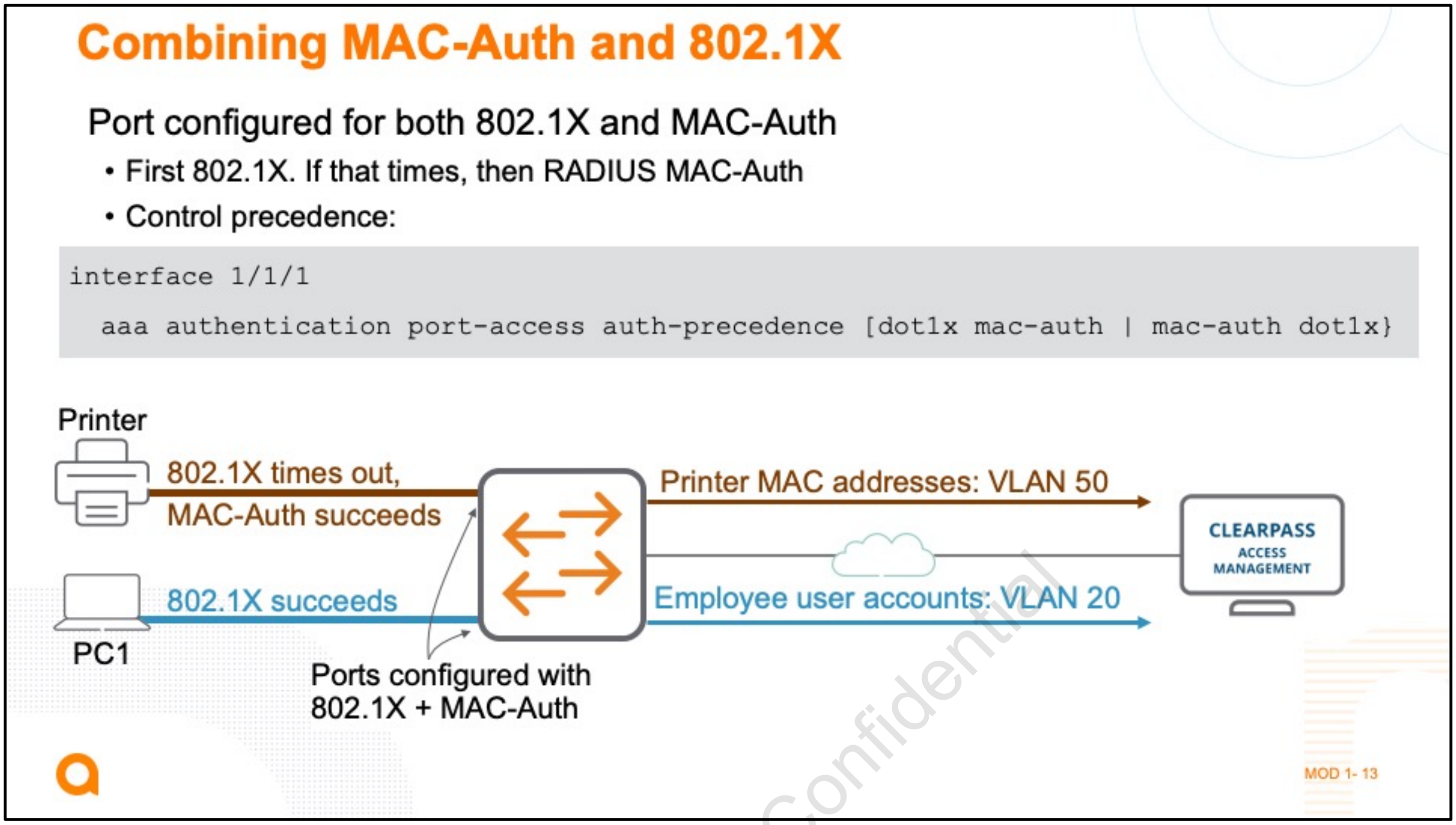
In another scenario, you might have a hub, an unmanaged switch, or an unmanaged AP to which multiple devices connect.

For these scenarios, set up LMA or RADIUS MAC-Auth as needed. By default, the switch allows only one MAC address to authenticate on the port at the same time. After a connected client passes authentication, the switch drops traffic received from any other MAC address even if that address would otherwise pass authentication. You must raise the MAC address limit as shown in the figure. You can set the value from 1 (default) to 256.

You typically assign tagged VLANs to IP phones and untagged VLANs to computers. If you want to enforce authentication for the computers, you can assign the untagged VLAN through the RADIUS profile, or you can have computers use the port's static VLAN.

Note: 802.1X port access and MAC authentication can be configured at the same time on a port. A total of 256 clients can be configured per port and 16,384 clients on the entire switch, irrespective of the authentication method. After the limit of 16,384 clients is reached, no additional authentication clients are allowed on any port for any method. The default is one client.

MAC authentication, MAC lockout, and port security are mutually exclusive on a given port. If you configure any of these authentication methods on a port, you must disable LACP on the port.



Combining MAC-Auth and 802.1X

Port configured for both 802.1X and MAC-Auth

- First 802.1X. If that times, then RADIUS MAC-Auth
- Control precedence:

```
interface 1/1/1
  aaa authentication port-access auth-precedence [dot1x mac-auth | mac-auth dot1x}
```

A single port can support multiple authentication methods, enabling it to authenticate different types of devices and users. For example, you have devices that use 802.1X authentication, but also some devices that are not. By default, 802.1X is performed first. If 802.1X times out, the switch assumes that the device doesn't support it. If MAC-Auth is also enabled on the port, then the switch will authenticate the device with MAC-Auth. You can change the order, per-interface, as shown in the figure.

In this scenario, each port supports a single device, but you are not sure which type of device will connect. You want any port to support any device type – those that support 802.1X and those that only support RADIUS MAC-Auth. Suppose you left it at the default, so 802.1X is tried first.

When the printer connects, the switch initiates 802.1X authentication, which the printer does not support and so does not respond. 802.1X times out, and the switch tries MAC-Auth, which succeeds. Policy defined on ClearPass ensures that printer ports are assigned to VLAN 50.

When the PC connects, the switch initiates 802.1X authentication, which the PC supports and so authentication succeeds. ClearPass policy ensures that employees are assigned to VLAN 20.

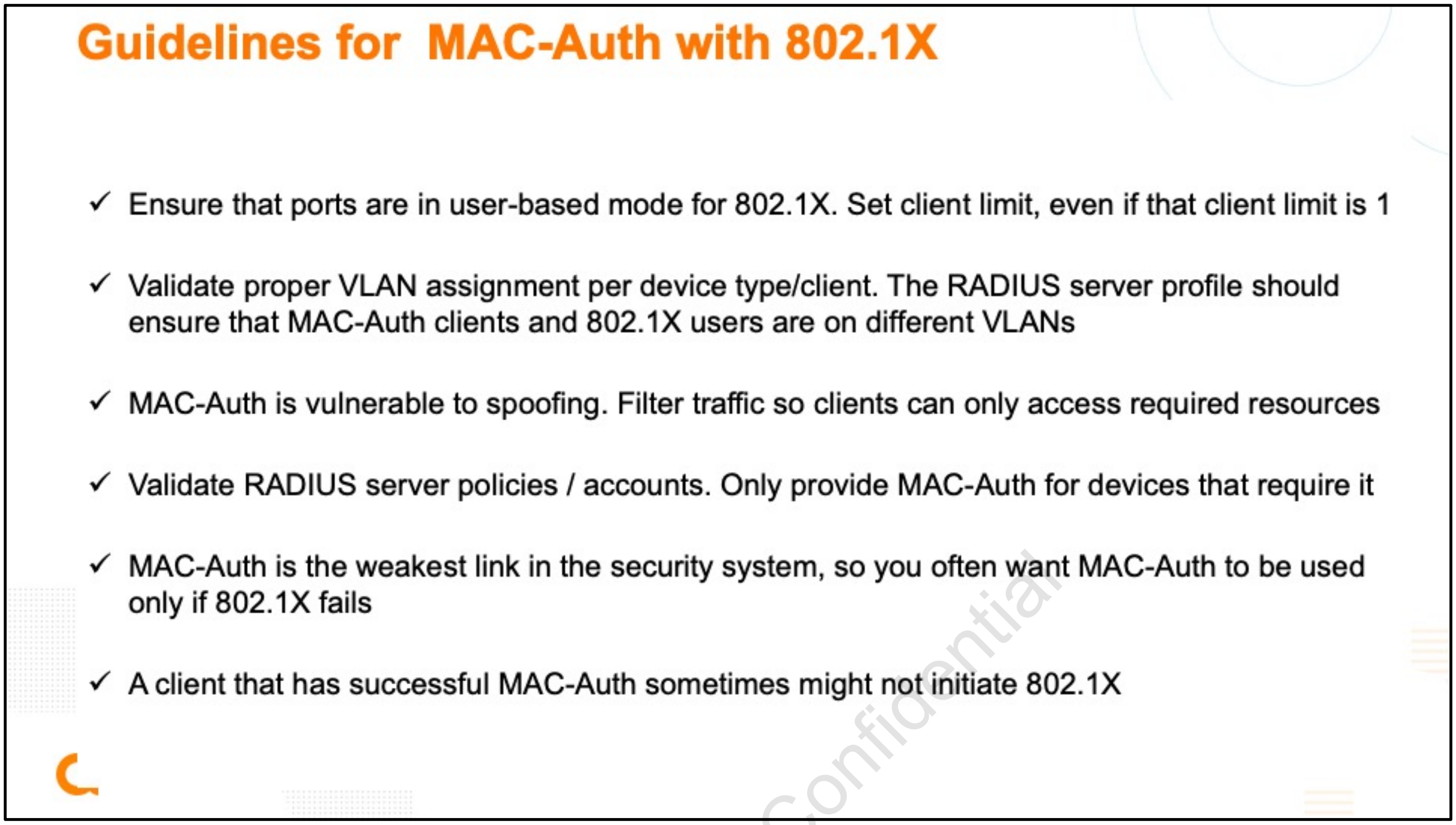
Suppose that you change the order to try MAC-Auth first. The port uses the device's MAC address to submit an access request, just as it would if MAC-Auth were the only authentication method:

If RADIUS MAC-Auth succeeds, the device is authorized for network access with the settings returned by the RADIUS server profile.

If RADIUS MAC-Auth fails, the user still has the option to pass 802.1X authentication. If the user succeeds, the device is authorized for network access with the settings returned by the RADIUS server.

If both types of authentication fail, the device is prohibited access.

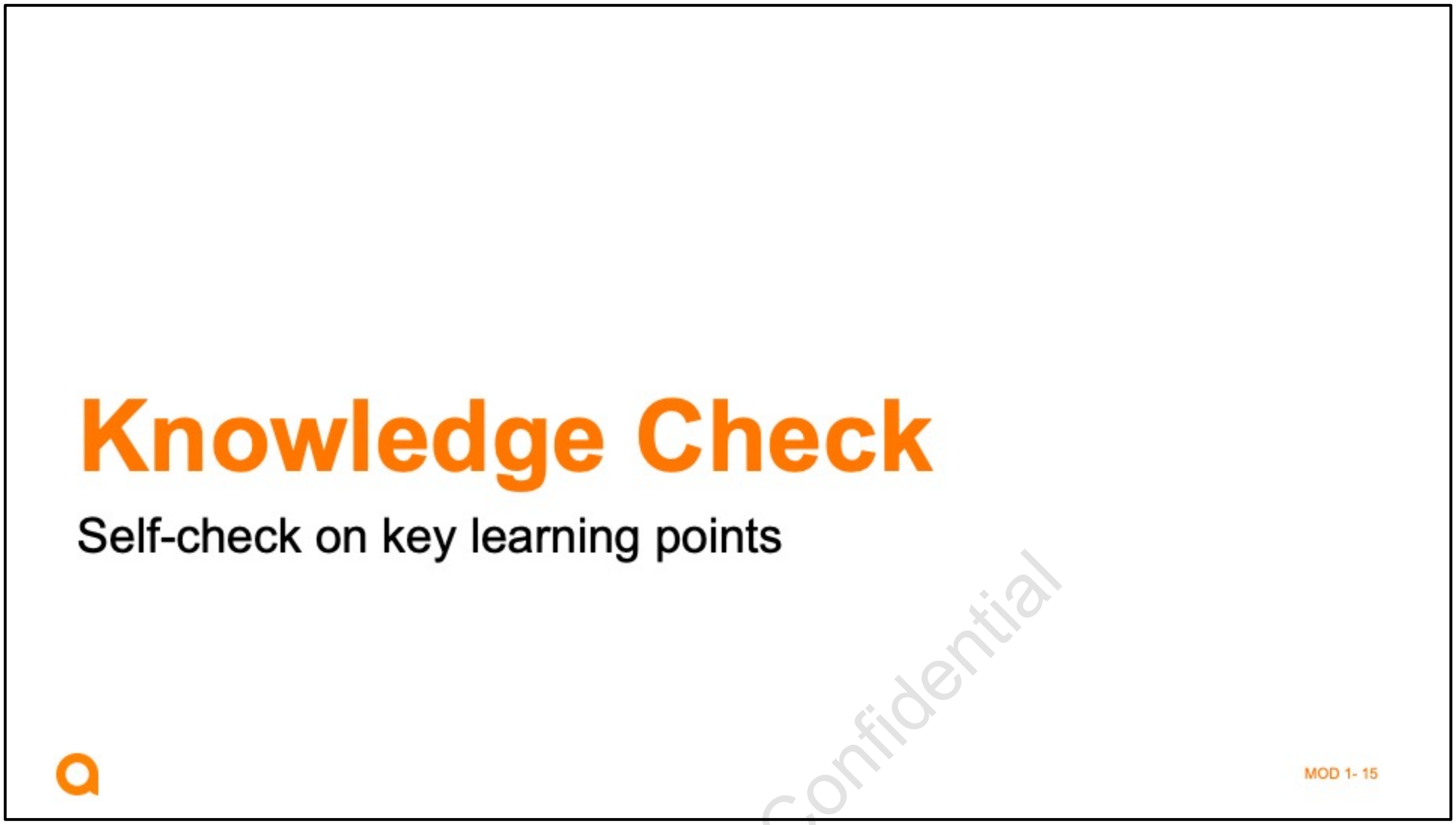
Even if RADIUS MAC-Auth succeeds, the user device might initiate 802.1X. (However, the switch itself will not.) If the user then authenticates successfully with 802.1X, the settings returned at the end of the 802.1X process take precedence over any settings applied by RADIUS MAC-Auth.



Guidelines for MAC-Auth with 802.1X

- ✓ Ensure that ports are in user-based mode for 802.1X. Set client limit, even if that client limit is 1
- ✓ Validate proper VLAN assignment per device type/client. The RADIUS server profile should ensure that MAC-Auth clients and 802.1X users are on different VLANs
- ✓ MAC-Auth is vulnerable to spoofing. Filter traffic so clients can only access required resources
- ✓ Validate RADIUS server policies / accounts. Only provide MAC-Auth for devices that require it
- ✓ MAC-Auth is the weakest link in the security system, so you often want MAC-Auth to be used only if 802.1X fails
- ✓ A client that has successful MAC-Auth sometimes might not initiate 802.1X

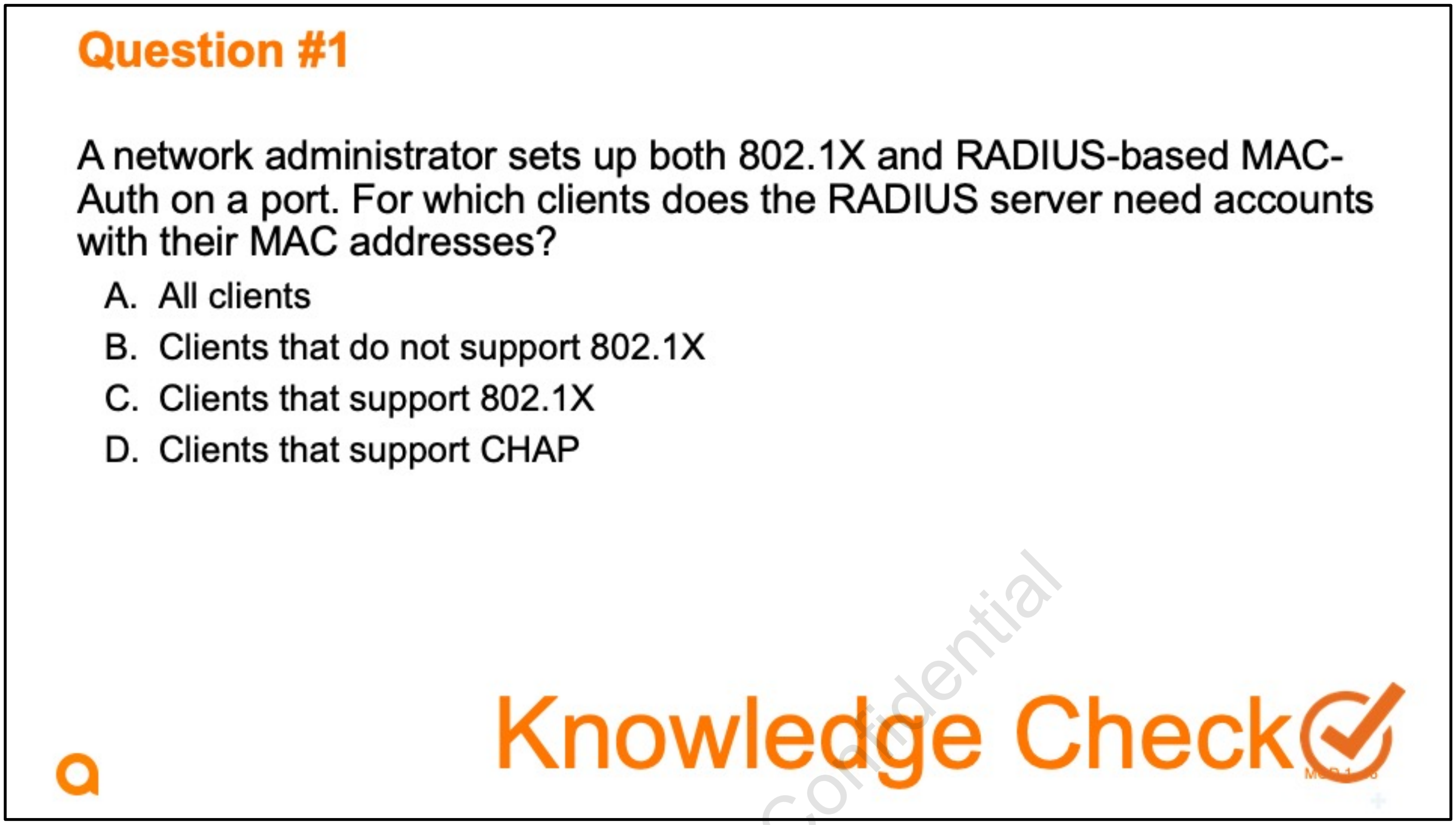
The figure highlights important guidelines when enabling MAC-Auth or LMA and 802.1X on the same port.



Knowledge Check

Self-check on key learning points

MOD 1- 15



Question #1

A network administrator sets up both 802.1X and RADIUS-based MAC-Auth on a port. For which clients does the RADIUS server need accounts with their MAC addresses?

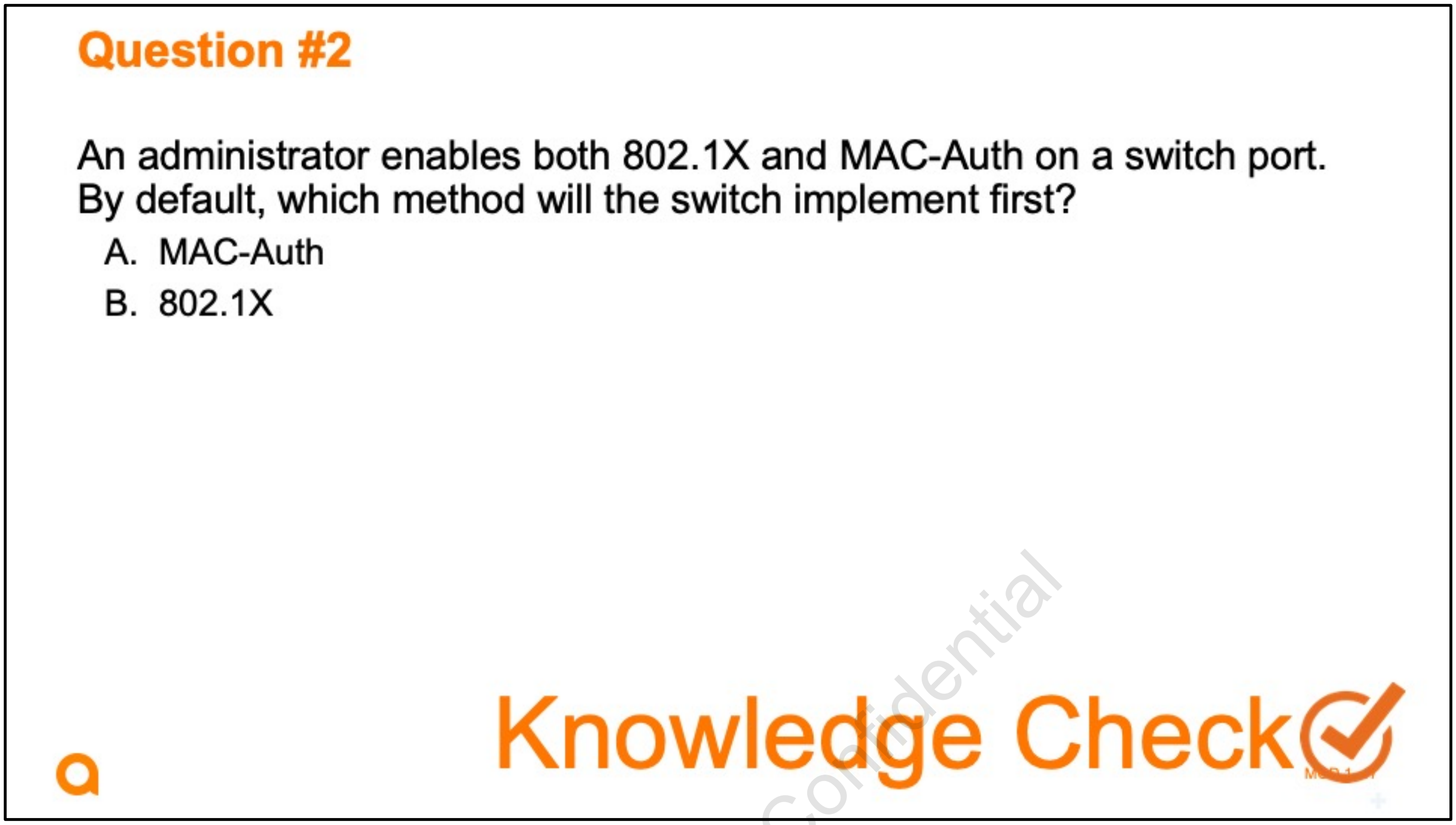
A. All clients

B. Clients that do not support 802.1X

C. Clients that support 802.1X

D. Clients that support CHAP

Knowledge Check



Question #2

An administrator enables both 802.1X and MAC-Auth on a switch port. By default, which method will the switch implement first?

A. MAC-Auth

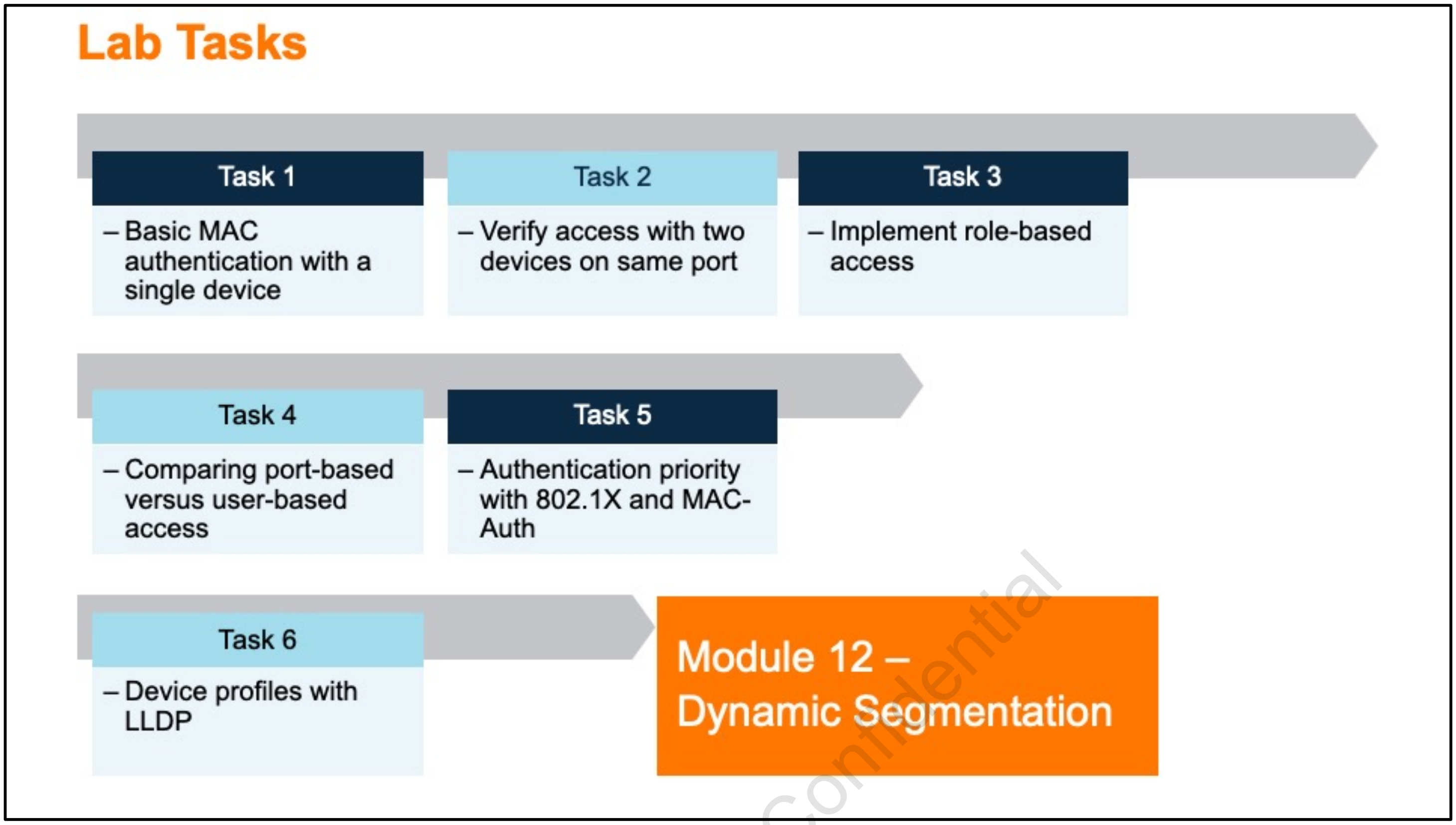
B. 802.1X

Knowledge Check

Lab Activity

Lab 6.3

Let’s do a lab.



Lab Tasks

Task 1

– Basic MAC authentication with a single device

Task 2

– Verify access with two devices on same port

Task 3

– Implement role-based access

Task 4

– Comparing port-based versus user-based access

Task 5

– Authentication priority with 802.1X and MAC-Auth

Task 6

– Device profiles with LLDP

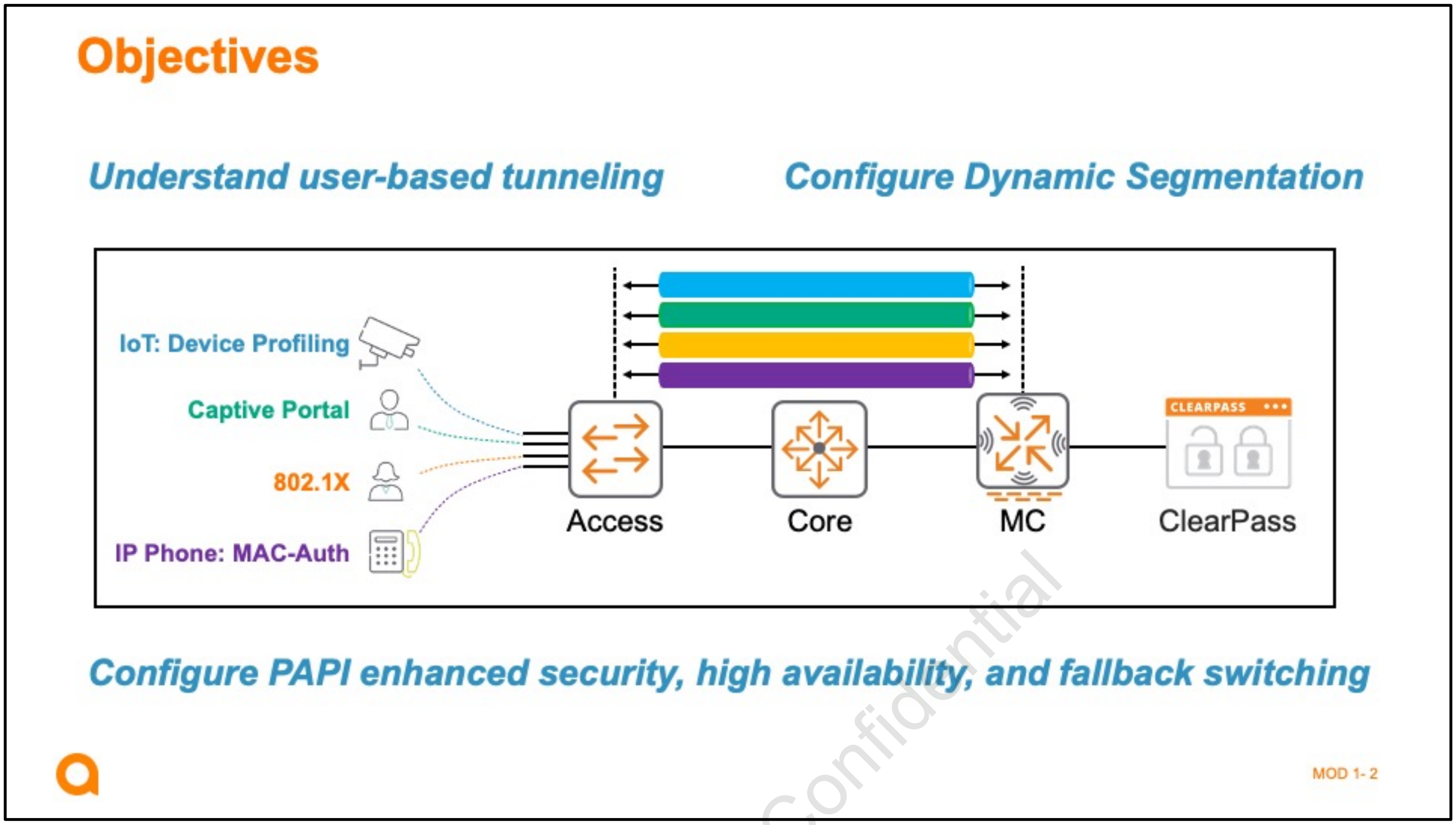
Module 12 – Dynamic Segmentation

The figure provides a brief review of lab tasks. Please see your lab guide for details.

Rev # 20.21

Dynamic Segmentation

Implementing ArubaOS-CX Switching



Objectives

Understand user-based tunneling

Configure Dynamic Segmentation

Configure PAPI enhanced security, high availability, and fallback switching

After completing this module, you should be able to:

Understanding the operation of user-based tunneling

Configure dynamic segmentation on AOS-CX switches

Describe when and how to configure PAPI enhanced security, high availability, and fallback switching for tunneled-node

Overview

- Overview
- UBT Configuration
- High Availability
- Verification
- Lab Activities

MOD 1- 3

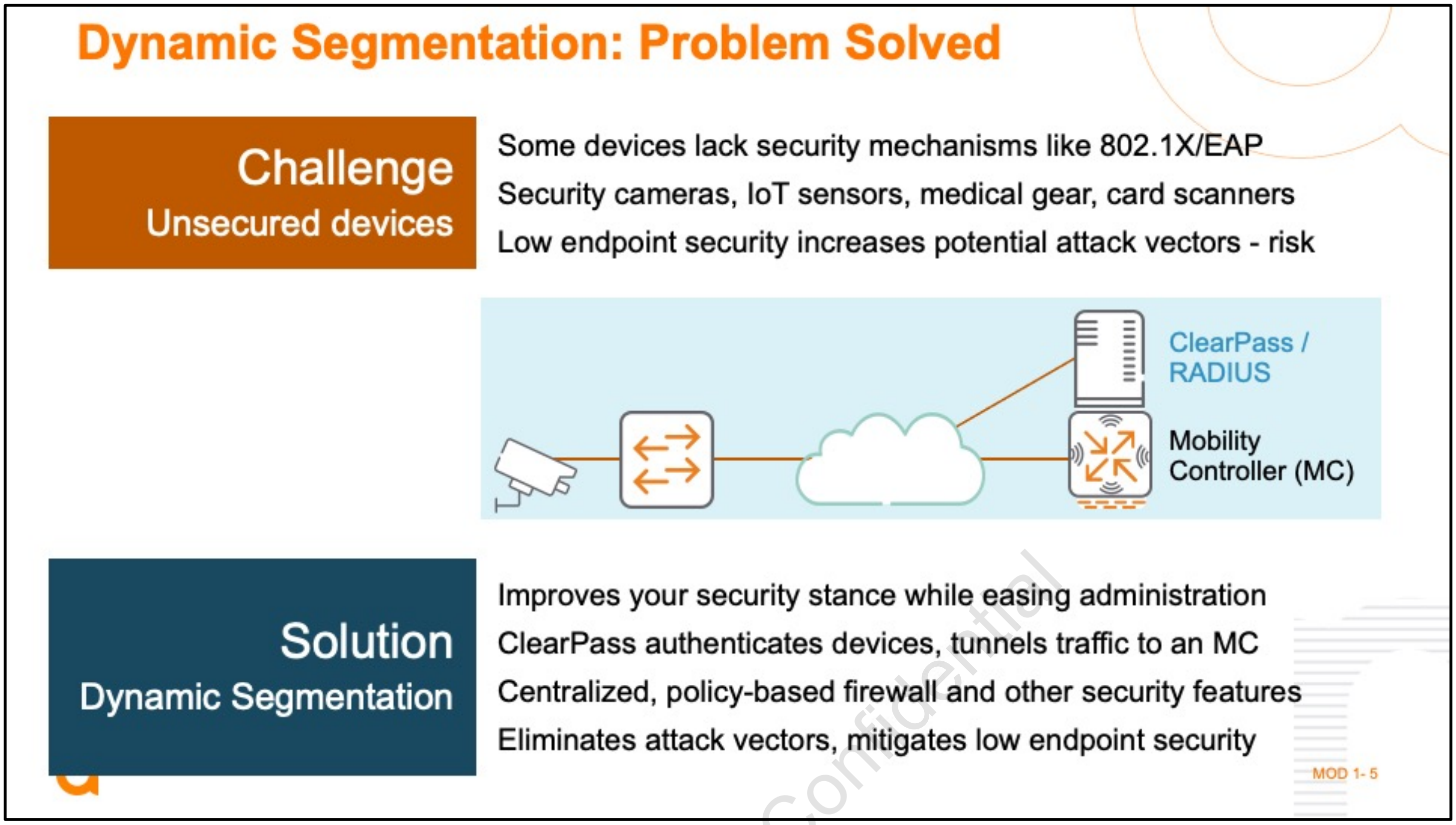
This course focuses on AOS-CX switches, but switches and Aruba Mobility Controllers (MCs) and APs will likely only be part of your solution. In addition to Aruba Central or AirWave (for network management) and Aruba ClearPass (for user access control), you may also have Aruba APs connected to Mobility Controllers (MCs). You might want to leverage the access control, firewall, and packet inspection capabilities of the MCs that currently secure your wireless devices to secure your wired endpoints as well.

In this module, you will learn how to use the dynamic segmentation feature, commonly called tunneled-node, for this purpose.

Overview

aruba
a Hewlett Packard Enterprise company

We begin with an overview.



Dynamic Segmentation: Problem Solved

Challenge
Unsecured devices

Some devices lack security mechanisms like 802.1X/EAP
Security cameras, IoT sensors, medical gear, card scanners
Low endpoint security increases potential attack vectors - risk

ClearPass /
RADIUS
Mobility
Controller (MC)

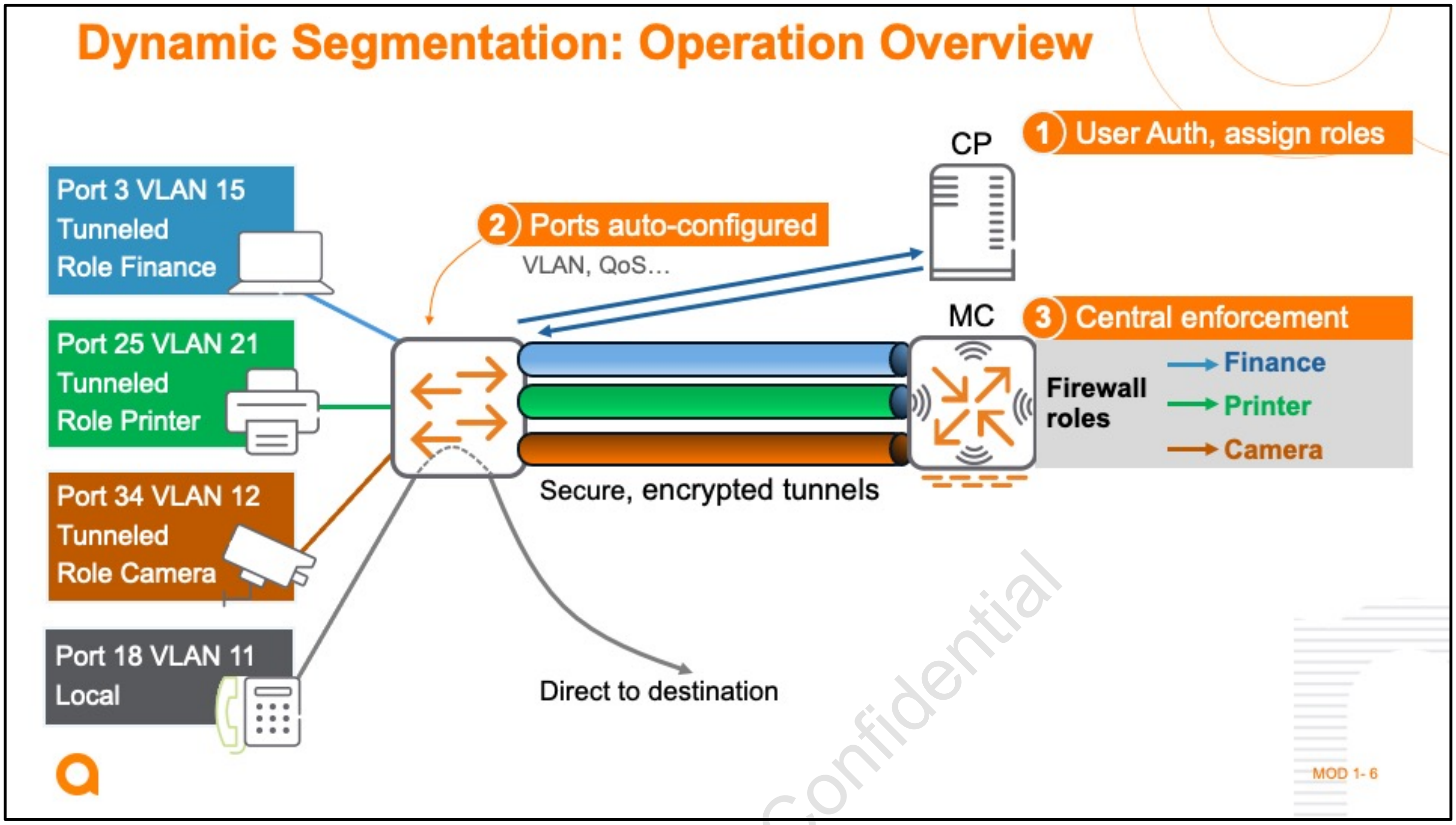
Solution
Dynamic Segmentation

Improves your security stance while easing administration
ClearPass authenticates devices, tunnels traffic to an MC
Centralized, policy-based firewall and other security features
Eliminates attack vectors, mitigates low endpoint security

MOD 1- 5

All modern desktops, laptops, tablets, and smart phones support robust security mechanisms like 802.1X/EAP. But many devices lack these security mechanisms. This includes specialty devices, and some devices that rely on Power over Ethernet (PoE), such as security cameras, payment card readers, Internet of Things (IoT), and medical devices. This lack of endpoint security can pose a serious risk to your infrastructure.

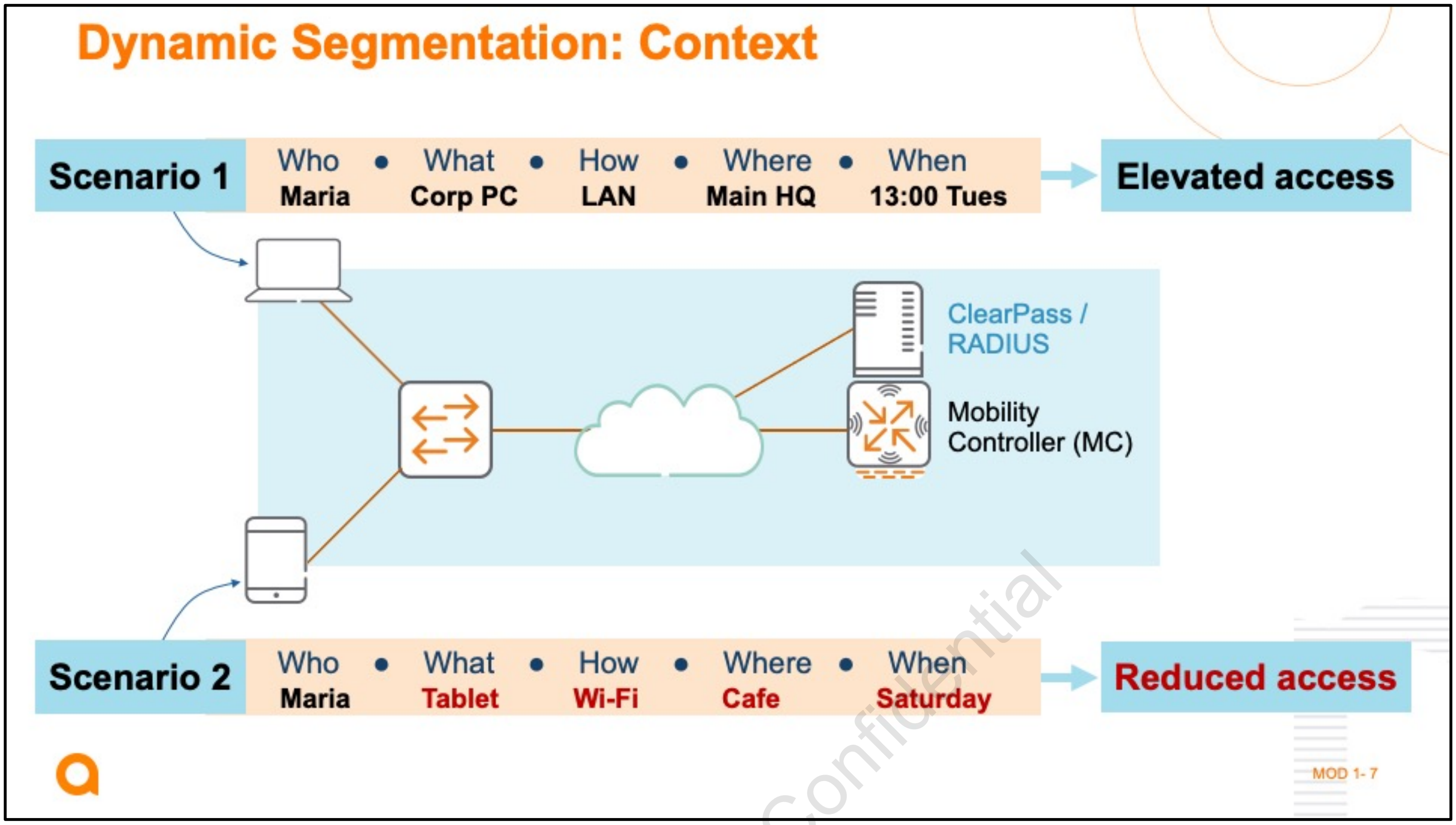
A switch with dynamic segmentation improves your security stance while easing administration. It authenticates these devices using ClearPass, and tunnels the traffic to a mobility controller, which centralizes formerly disparate Access Lists (ACLs) and firewall rules into single, intuitive policy system. You can continue to use these devices without compromising network security.



Dynamic Segmentation: Operation Overview

ClearPass profiles each device, deems appropriate endpoints as “acceptable”, and sends accept messages to the controller. The messages include tunnel information, assigned VLAN, and secondary role. In the figure, a user was profiled into the Finance role, and so their port is configured for VLAN 15, and their traffic is securely tunneled to the controller. The controller applies centralized policy enforcement to limit finance traffic as appropriate. Printers and cameras are similarly treated, as shown in the figure.

ClearPass can also indicate if device traffic should be locally forwarded by the switch, directly toward its ultimate destination, without tunneling to the controller. You sacrifice some of that centralized PEF enforcement, but you accommodate devices that are uber-sensitive to delay and jitter. You just don’t get all the benefits of this “tunneled-node” solution.



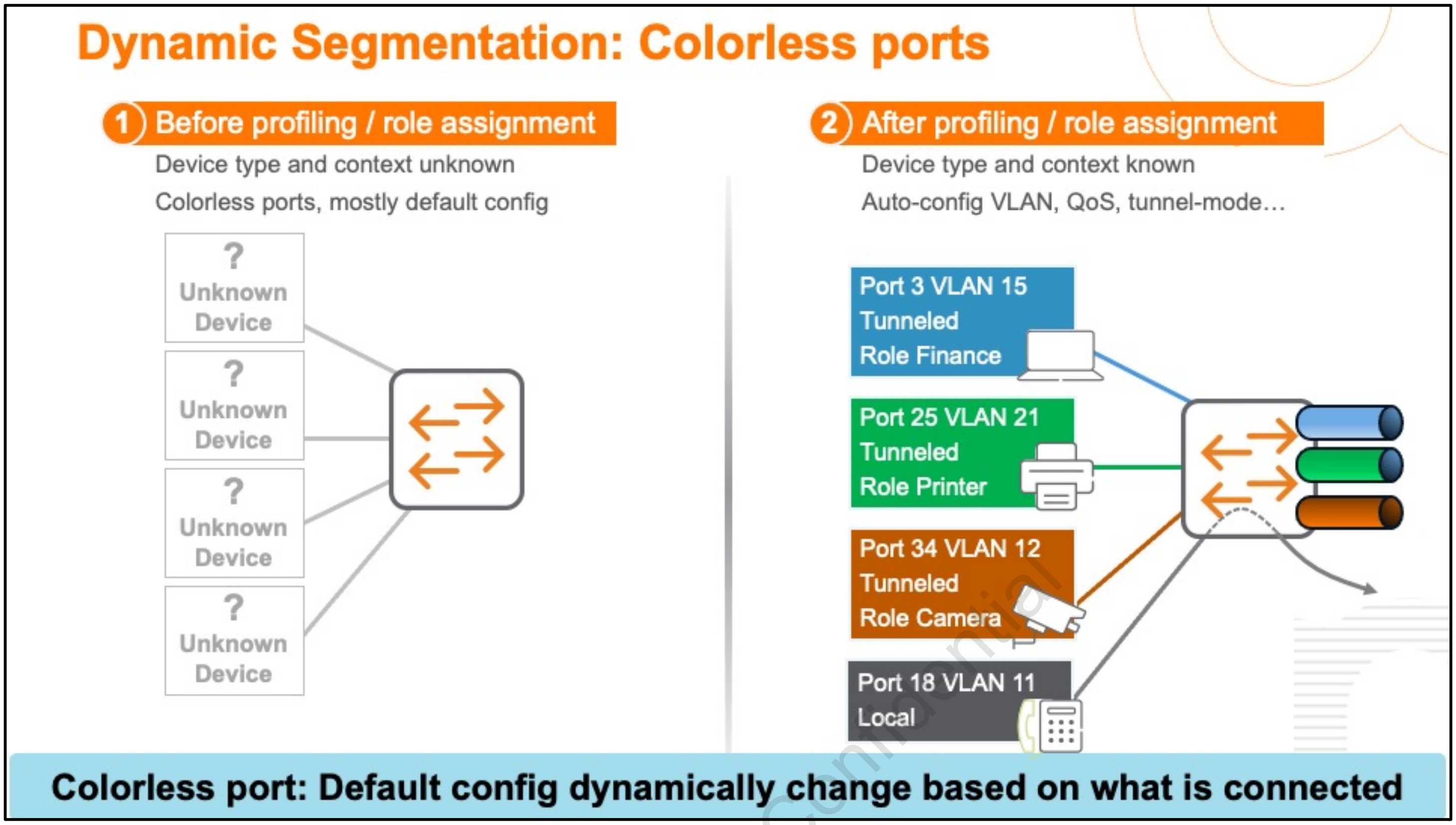
Dynamic Segmentation: Context

	Who	What	How	Where	When	
Scenario 1	Maria	Corp PC	LAN	Main HQ	13:00 Tues	Elevated access
Scenario 2	Maria	Tablet	Wi-Fi	Cafe	Saturday	Reduced access

MOD 1- 7

ClearPass can leverage device profiling to auto-determine device types, and a rich set of contextual “who-what-how-where-when” information to assign policies. Maria’s login credentials might give her elevated access when using her corporate PC connected via wireless at the main building, during business hours. But when she uses her personal tablet connected from a local café on Saturday, she may have reduced access.

Among other things, this helps you to implement colorless ports on your switches.



Dynamic Segmentation: Colorless ports

1) Before profiling / role assignment

Device type and context unknown

Colorless ports, mostly default config

2) After profiling / role assignment

Device type and context known

Auto-config VLAN, QoS, tunnel-mode…

Port 3 VLAN 15 Tunneled Role Finance

Port 25 VLAN 21 Tunneled Role Printer

Port 34 VLAN 12 Tunneled Role Camera

Port 18 VLAN 11 Local

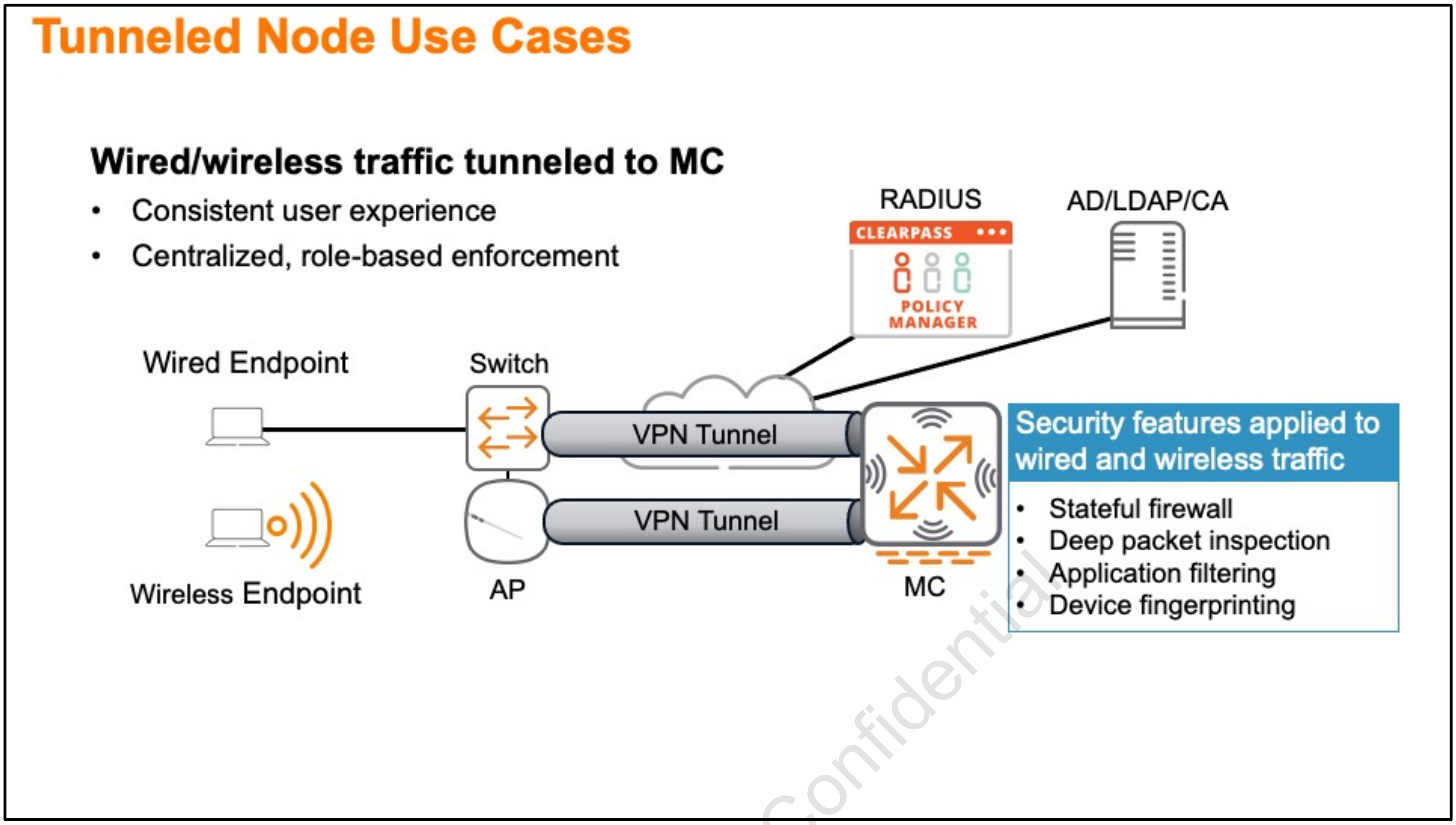
Colorless port: Default config dynamically change based on what is connected

A colorless port is basically a port with default characteristics that are then dynamically changed based on device characteristics and authentication details. Look at the left-side figure. Before profiling, context assessment, and role assignment the device type and context is largely unknown.

Manually configuring VLANs, QoS, ACLs, and more is tedious, time-consuming, and error prone – and you are now free from much of this!

Leave the ports at default settings. Connected devices are automatically profiled context is assessed, and roles are assigned. Today the finance person connected to port 3 on the switch. Tomorrow they may connect to port 12. It doesn't matter – dynamic segmentation "colors" the port automatically.

Colorless Ports at the most basic level, can be demonstrated with a single standalone switch using local user roles or device profiling with LLDP. Without using external RADIUS or tunneling, a customer can observe how the same port takes on different policies (QoS/ACL/VLANs) depending on what device connects to the port.



Tunneled Node Use Cases

Wired/wireless traffic tunneled to MC

- Consistent user experience
- Centralized, role-based enforcement

Security features applied to wired and wireless traffic

- Stateful firewall
- Deep packet inspection
- Application filtering
- Device fingerprinting

Aruba ClearPass and port-based authentication implemented on switches provide one method for unifying access policies. Both the MCs and AOS-CX switches can enforce 802.1X authentication to users and apply the policies returned by ClearPass. However, ClearPass cannot simply send the same settings to both MCs and switches. For example, ClearPass often simply sends a user role assigned to an MC, and the MC has firewall rules associated with that role. On the other hand, ClearPass would need to send a policy with dynamic ACLs to apply similar controls to an AOS-CX switch.

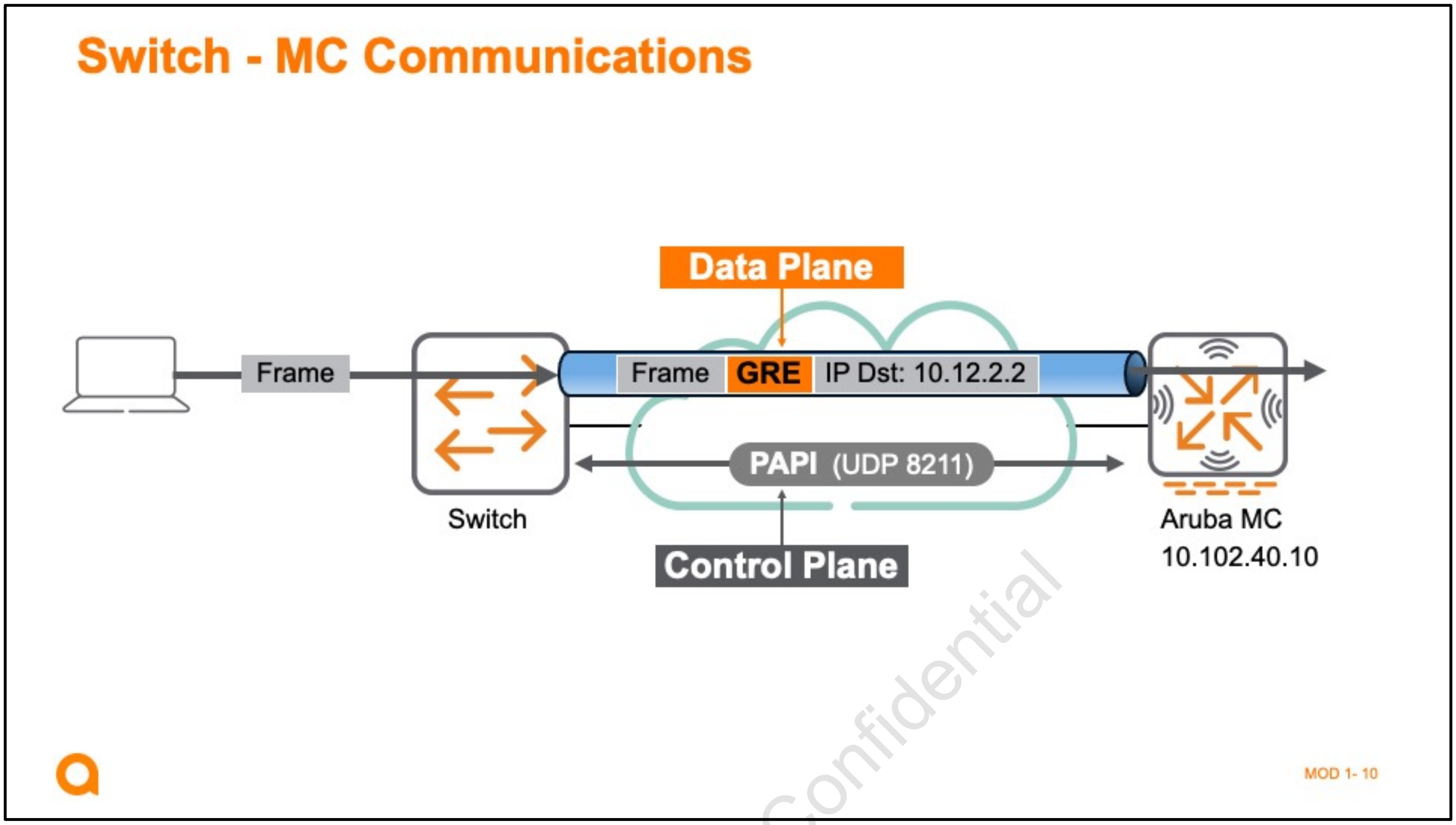
Tunneled-node provides tighter and simpler unification of wired and wireless access. Switches simply tunnel traffic to the MC, which handles the traffic much as it would handle traffic tunneled from an AP. An employee can connect through their desktop, log in, be assigned a role, and receive customized access. Later the employee can connect wirelessly using a laptop in a meeting room, log in, be assigned the same role, and receive the same access.

Do keep in mind that tunneling the wired devices’ traffic to the MC, of course, increases the demands on that device. Architects will need to plan the Aruba solution’s capacity and licensing accordingly. A list of general guidelines is provided later in this module.

Note: Dynamic segmentation is supported on both Aruba switching products: AOS-CX and

AOS switches.

The bottom line is this: Use the tunneled-node feature whenever you want to apply Aruba controller-based security/control mechanisms to both wireless and wired traffic. You get unified access control – same policies regardless of whether they have a wired or wireless connection.

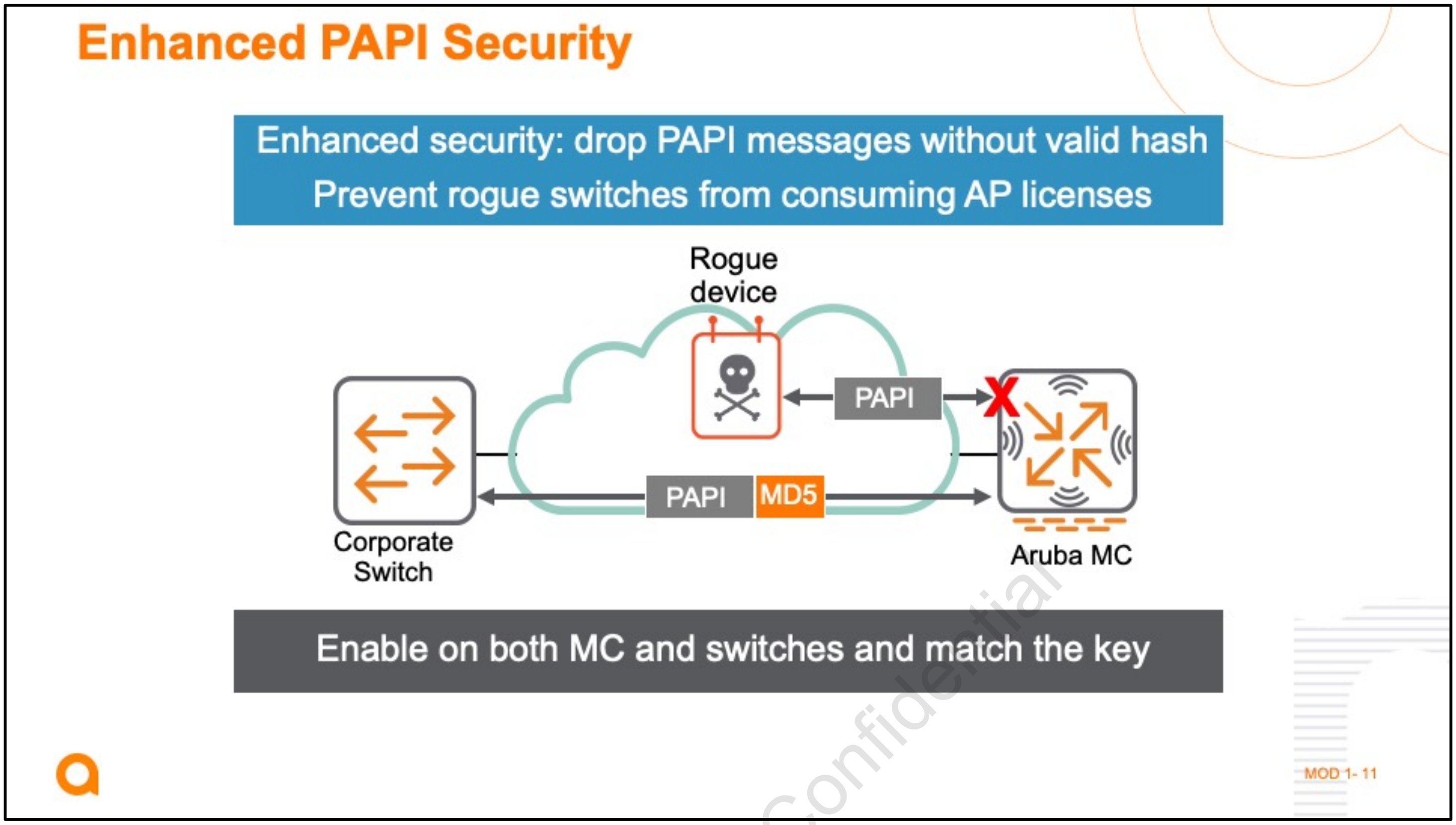


Switch - MC Communications

The switch uses two protocols to connect to an Aruba Mobility Controller (MC)

The control plane uses PAPI (UDP port 8211) - the same protocol used by AP-to-MC communications. However, where APs use IPSec to protect the PAPI connection between the AP and MC, AOS-CX switches do not support this protection. Instead, you can optionally implement an MD5 HMAC function to protect PAPI between the AOS-CX switches and MCs.

The data plane uses Generic Routing Encapsulation (GRE). Endpoints send standard frames to the switch, which encapsulates them in GRE and tunnels them to the MC.



Enhanced PAPI Security

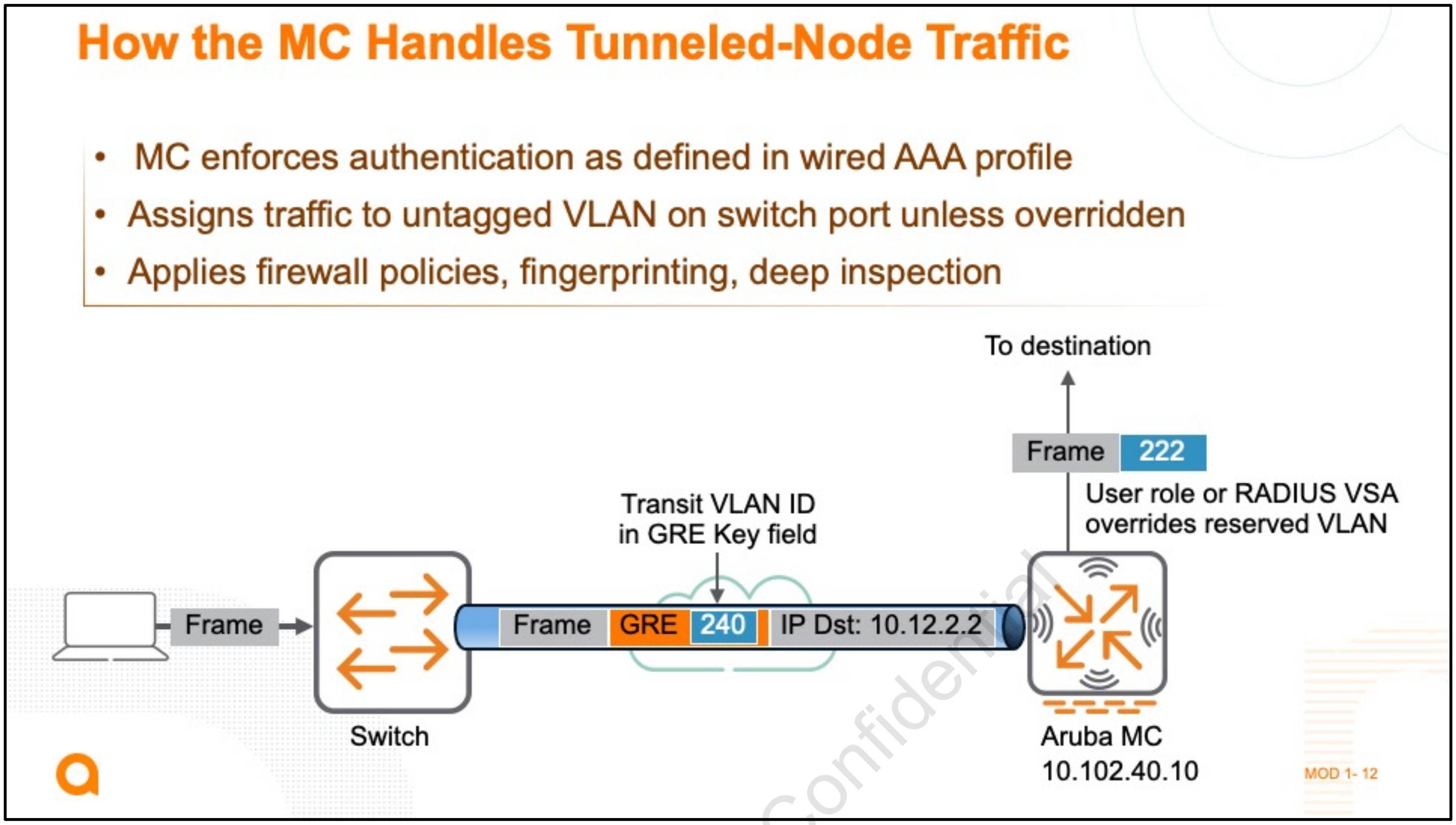
Enhanced security: drop PAPI messages without valid hash
Prevent rogue switches from consuming AP licenses

Enable on both MC and switches and match the key

Switch-MC PAPI control plane traffic creates a potential security risk. Rogue devices could send PAPI messages to APs, MCs, and switches and disrupt functionality. They could also try to insert spoofed PAPI messages on the controller to disrupt communications or obtain information about the network.

PAPI enhanced security minimizes this risk by using a MD5 and user-configurable key to hash the message. If a rogue device tries to spoof or tamper with a message, the hash will not check out because the rogue does not know the correct key to create the hash. A primary reason for this enhanced security is to prevent rogue switches from consuming AP licenses on the controller.

You must configure matching keys on the MC and switch - between 10 and 64 characters. Enhanced security mode is disabled by default.



How the MC Handles Tunneled-Node Traffic

- MC enforces authentication as defined in wired AAA profile
- Assigns traffic to untagged VLAN on switch port unless overridden
- Applies firewall policies, fingerprinting, deep inspection

The MC divides the arriving tunneled traffic into one tunnel per tunneled-node port. It decapsulates the traffic and applies various controls. Learning how to configure the MC to control tunneled-node traffic is beyond the scope of this course, but you should understand how the configuration affects the MC's decisions.

The switch communicates the untagged VLAN configured on a tunneled-node port to the MC. The MC must have this VLAN configured on it. It then handles the tunneled traffic based on the settings in the wired AAA profile assigned to that VLAN. For example, the MC might apply MAC Auth or 802.1X—or some combination. After successful authentication, the controller applies a role to the traffic. Based on that role, it controls traffic with firewall policies and other policy actions. Finally, it forwards the packet towards its destination.

You must also set up the MC to forward the traffic in the tunneled-node ports' untagged VLAN at Layer 2 or Layer 3, as desired. In the figure, the MC uses Layer 2 forwarding - it forwards the frame tagged for VLAN 240. If the MC were using Layer 3 forwarding, it would have an IP address on VLAN 240 and route traffic out of this VLAN.

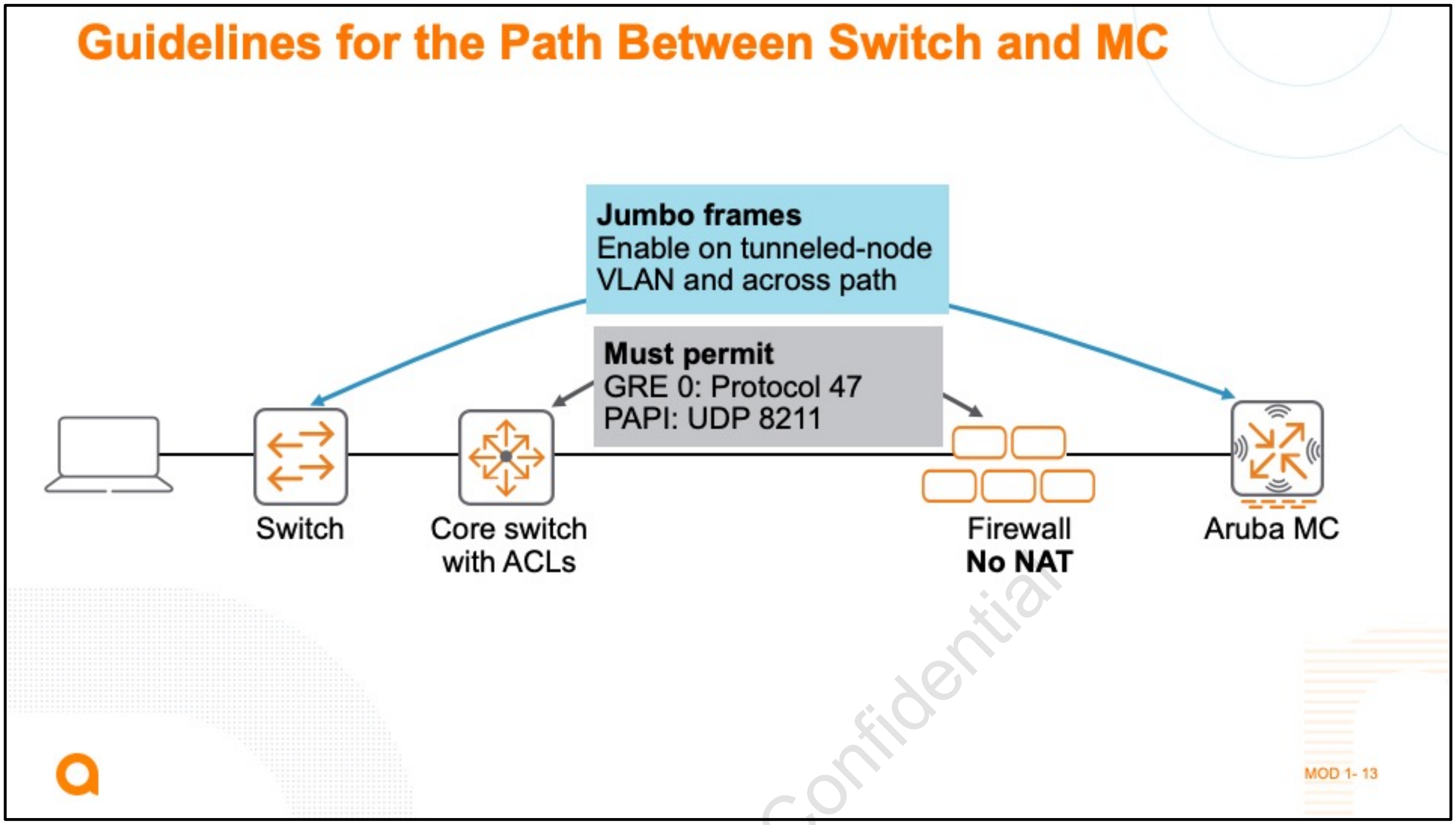
You can configure different VLANs on different tunneled-node ports to inform the controller to assign those nodes to different VLANs. However, the MC can also apply different firewall rules to different devices in the same VLAN, again based on dynamic settings applied during authentication. For example, the MC could enforce MAC-Auth and use the devices' MAC addresses to assign different roles to different devices. You should coordinate with MC administrators and make sure to match the switch configuration with their desired configuration.

Keep in mind that the switch is simply using the untagged VLAN on the tunneled-node port to communicate the desired default VLAN to the controller. It is not switching the tunneled-nodes' traffic in this VLAN, so you don't need to tag it on the switches' uplinks as you normally would. (The use of fallback switching mode makes an exception as you will learn later.) Instead the VLAN exists on the MC at the distribution or core. In keeping with this design, the switch is restricted from implementing many normal features—such as an IP address—on any untagged VLAN on a tunneled-node port. Aruba recommends reserving the tunneled-node VLANs for tunneled-node ports only.

You should also understand that the MC can override the default VLAN assignment after it authenticates the device. For example, it might apply 802.1X authentication to tunneled-nodes and receive a dynamic VLAN assignment from Aruba ClearPass. The MC would then forward the tunneled-node device traffic in this VLAN, rather than the untagged VLAN configured on the switch port. Even if you plan to apply dynamic VLANs for some devices, though, the MC must support the untagged VLAN on tunneled-node ports; the VLAN defines the wired AAA profile.

Finally, note that the figure above shows a single tunneled-node port on the switch for simplicity. When you apply tunneled-node to multiple ports, the MC sees one tunnel per port.

The term "transit" or "reserved" VLAN is used for the number placed in the GRE Header's Key field. This helps the MC determine the profile policies that should be used, by default, if nothing else is received, like the security policies and VLAN that should be applied to the user. In this example, the AAA server did not send a VLAN attribute, so what was used in the VLAN header is the final VLAN for the user on the MC. Note that this VLAN must be pre-configured on the MC. If the AAA server had passed down a specific VLAN attribute, like VLAN 888, then this would have become the user's final VLAN number and what the MC would use to tag the user's traffic before forwarding it onward.



Guidelines for the Path Between Switch and MC

Jumbo frames
Enable on tunneled-node VLAN and across path

Must permit
GRE 0: Protocol 47
PAPI: UDP 8211

MOD 1- 13

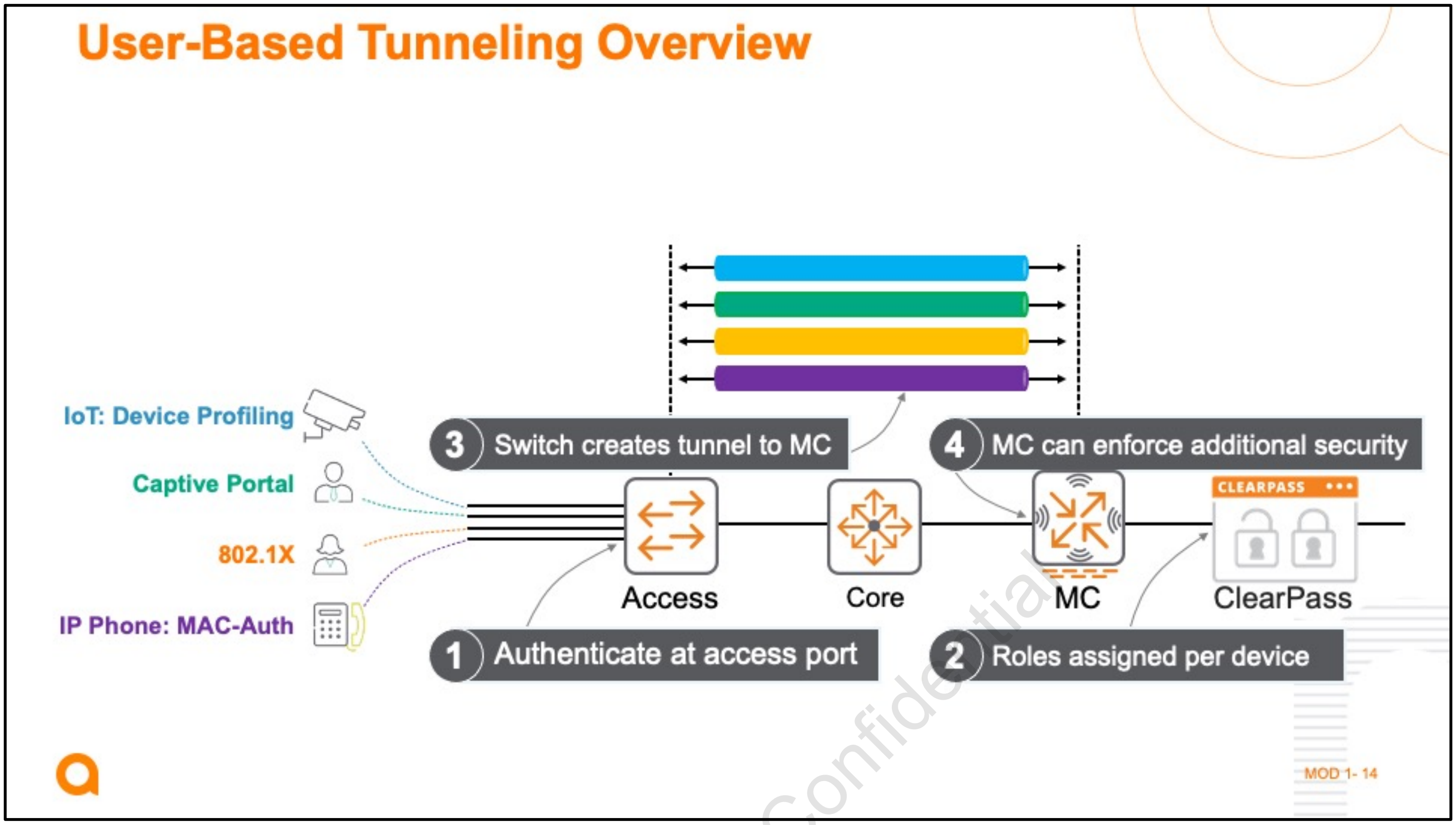
You should check the network infrastructure between the AOS-CX switches that implement tunneled-node and the MCs. Make sure that the infrastructure will not interfere with communications in any way.

If any firewalls or network infrastructure devices with ACLs are in the path, they must allow GRE and PAPI traffic. Enable GRE on IP protocol 47 and PAPI on UDP port 8211.

You must also understand that NAT cannot be applied in the path between the switch and the controller. It is best if you can direct the tunneled-node switch to a local MC.

Also keep in mind that encapsulating frames can cause them to exceed the MTU. However, tunneled-node ports do not support the features typically used to prevent MTU violations, including fragmentation, PMTU detection, or reassembly. Aruba recommends that you enable jumbo frames across the path from the switch to the MC. On AOS-CX switches, you enable jumbo frames on a VLAN basis. Make sure that every VLAN that carries traffic between the switch and the MC supports jumbo frames.

Make sure that you know the IP address that your AOS-CX switches will use to establish the tunnel with the MC. Typically, the switch acts at Layer 2 and has a single IP address and a default gateway for reaching all other subnets. But If your switch has multiple IP addresses, it uses the IP address on the forwarding VLAN for its route to the MC. Make sure that the MC can also reach this IP address (through its own route or its default gateway).



User-Based Tunneling Overview

1 Authenticate at access port

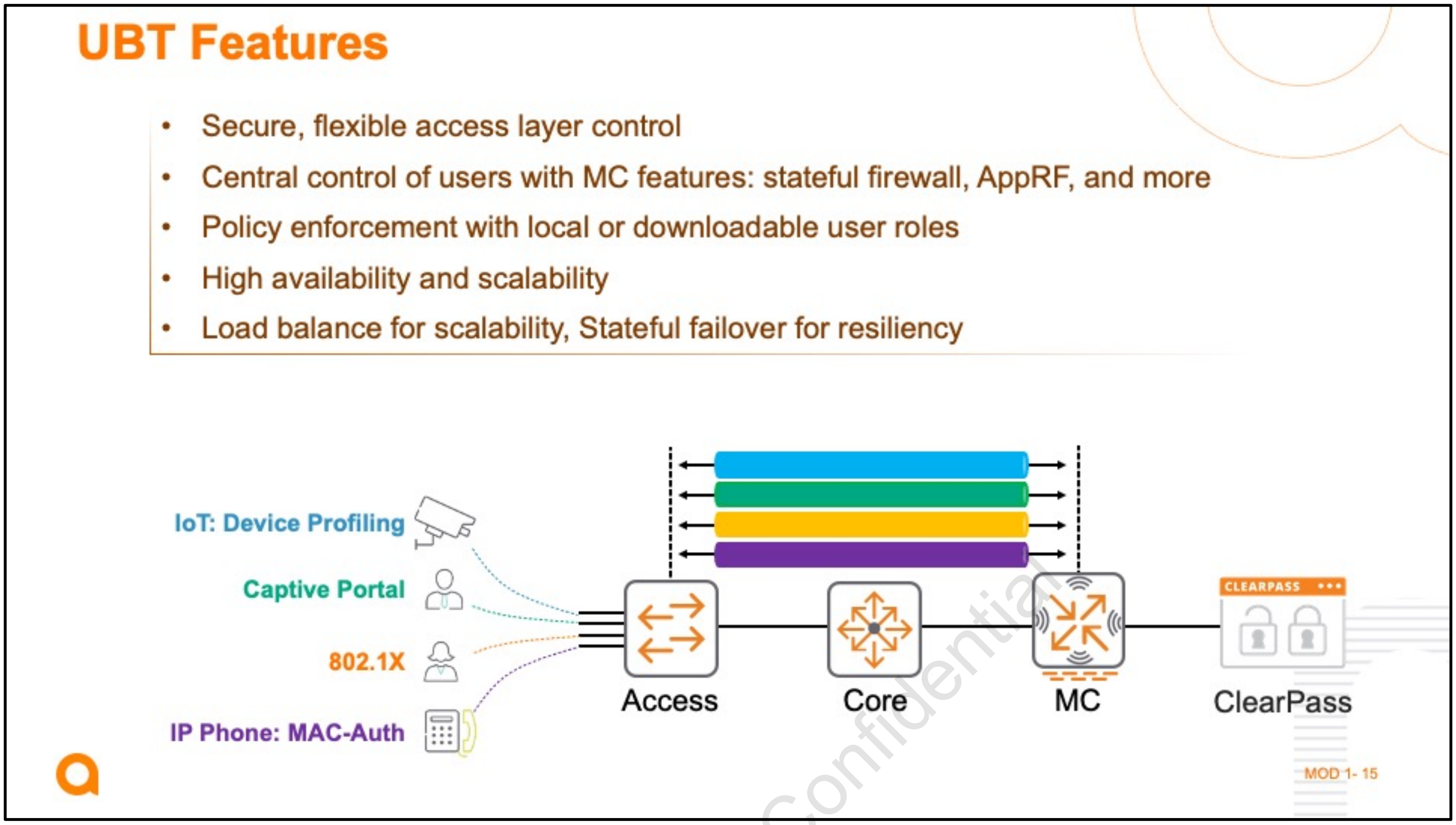
2 Roles assigned per device

3 Switch creates tunnel to MC

4 MC can enforce additional security

User Based Tunneling (UBT) uses the concept of a colorless access port. It doesn’t matter what you connect to the port; roles and policies are assigned per device.

As shown in the figure, some device connects to an access port – maybe an IoT device. Authentication happens at the port level, and ClearPass assigns a role. This role dictates how to color the port with VLAN, QoS, ACL, etc. The switch tunnels the end user traffic to the Mobility Controller, which can enforce additional security. Perhaps later the IoT device is disconnected, and a corporate user connects or IP phone connects to that same port. The port is colored, tunneled, and secured as appropriate.



UBT Features

- Secure, flexible access layer control
- Central control of users with MC features: stateful firewall, AppRF, and more
- Policy enforcement with local or downloadable user roles
- High availability and scalability
- Load balance for scalability, Stateful failover for resiliency

The figure summarizes key UBT features, as described below

Secured and flexible control of access layer: With ClearPass or switch configuration, only traffic from a specific user/device role is sent to the Mobility Controller Policies (e.g., QoS, ACL, rate-limit) can be enforced at tunneled ports or at the controller

Access to Controller’s applications: Users can access Controller’s applications such as stateful firewall and AppRF

Policy enforcement is achieved by local user roles or downloadable user roles

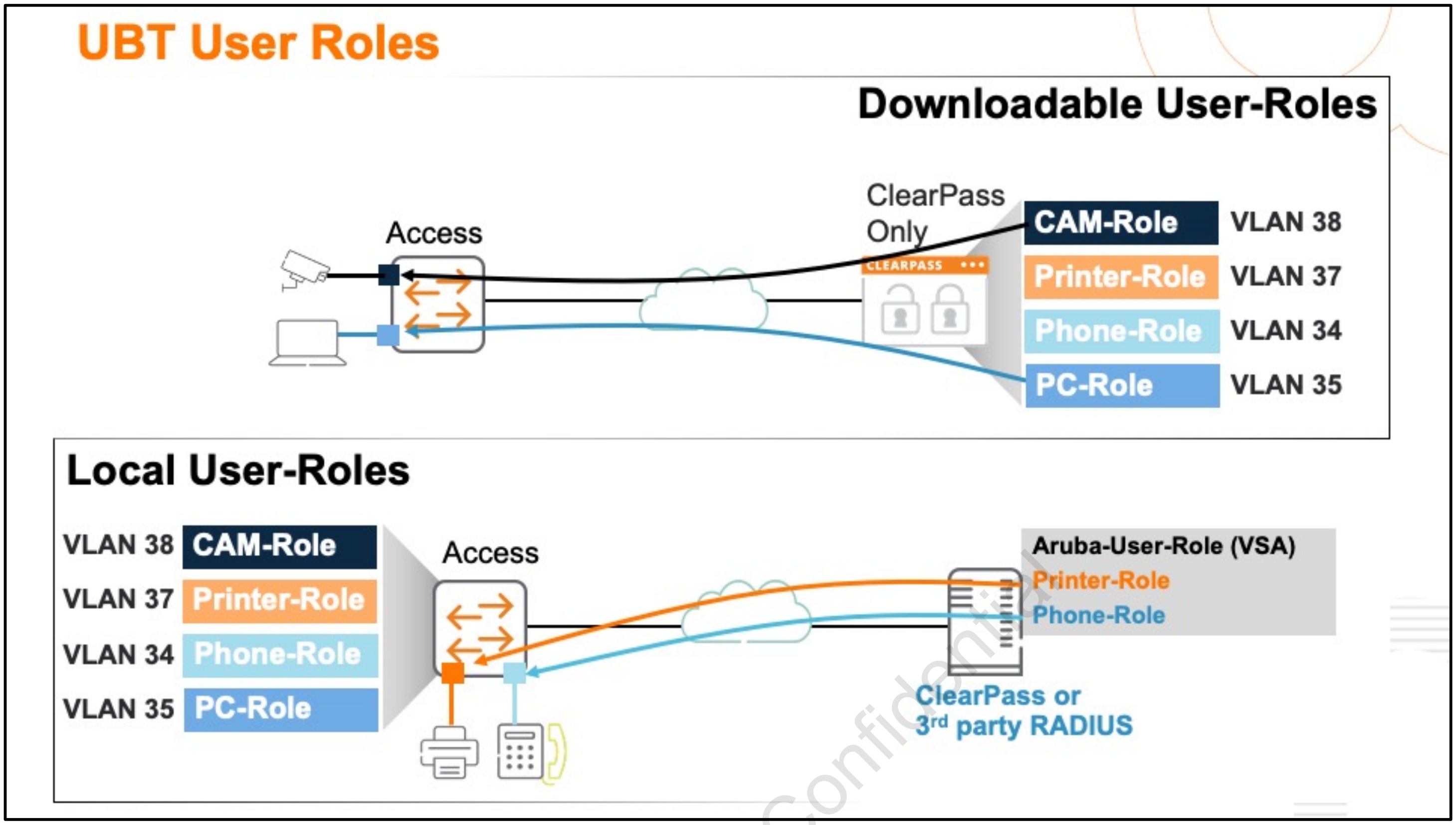
Local user roles are configured on the switch

Downloadable user roles are configured on ClearPass and pushed to the switch

High availability and scalability

Load balance to multiple controllers for high scalability

Stateful failover to standby mobility controller



There are two user roles that the switch can use:

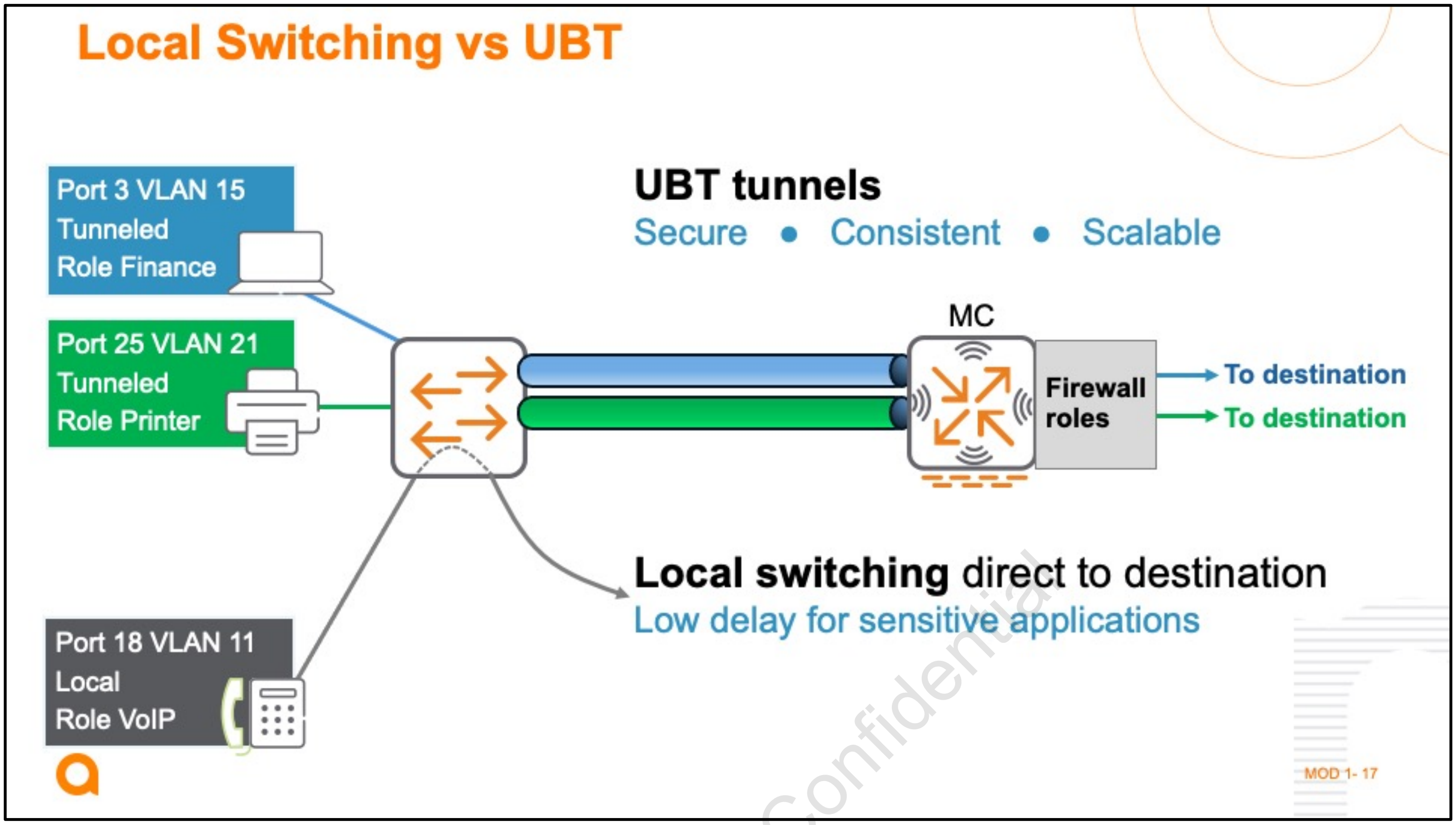
Local User Role (LUR): You define user roles locally on the switch. First some device connects and authenticates. Then the RADIUS server tells the switch which of its LUR's to apply by sending a User-Role name VSA. You can use ClearPass or a a third-party AAA server with LUR.

Downloadable User Role (DUR): You do not define roles on the switch. Instead you define them centrally on a ClearPass server – no third-party RADIUS support. When a device connects and authenticates successfully, the switch downloads the User-Role to the switch and the switch applies it to the device.

Note: On the older AOS switches, the term secondary role was used instead of gateway role. With AOS-CX switches, the appropriate term used to describe the role the switch passes to the controller is the gateway role.

Note: With AOS-CX release (10.4), Aruba does not support role download for a gateway (controller) or secondary role to the mobility controller. Either the role must pre-exist on the MC, or the user goes through a second authentication and the role can then be downloaded to the MC.

That's a review of the two methods used to assign user roles. Now let's look at the two methods switches can use to forward traffic.



Local Switching vs UBT

UBT tunnels

Secure • Consistent • Scalable

Local switching direct to destination

Low delay for sensitive applications

Local switching is typically used when delay-sensitive traffic is involved between access-layer devices, like voice or video communications (VoIP phones, for example) or a third-party firewall already exists in the network and the company wants to continue using the policy function of that firewall.

However, local switching lacks the centralized policy and enhanced security of UBT. Traffic filtering is limited to ACLs, with no stateful or application inspection features. In many cases, companies use only VLANs for security at the access layer, where devices in a VLAN can freely interact with each other. Their traffic is only processed by the firewall if it leaves the subnet. Allowing this type of promiscuous behavior can lead to many security issues - user devices can attack other user devices in the same subnet.

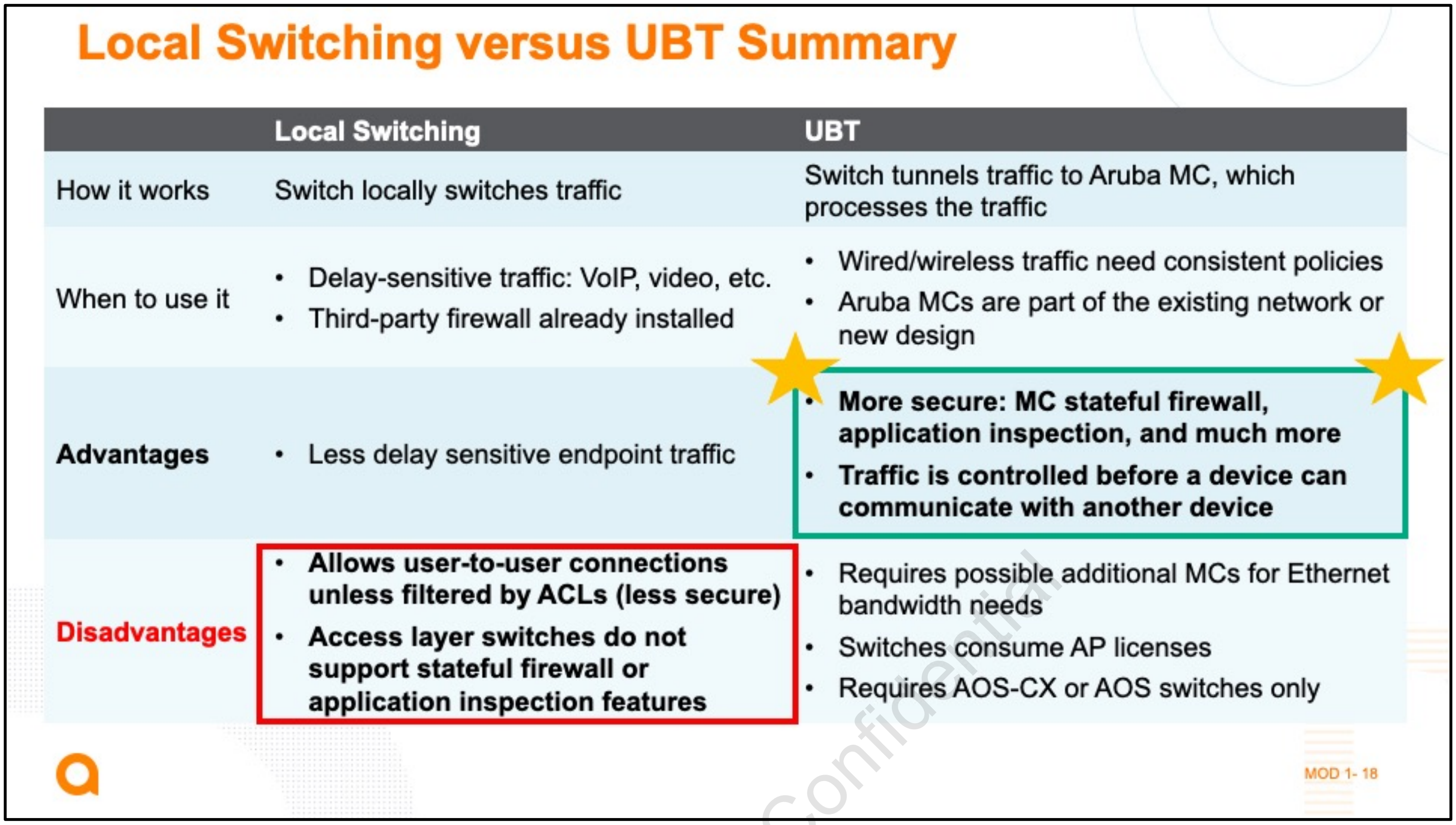
With UBT, the switch tunnels authenticated user traffic to an Aruba MC, to be processed by security policies. Advantages over local switching include:

Centralized security policies for both wired and/or wired traffic: users have a consistent experience whether they connect via wired Ethernet or Wi-Fi.

Endpoint traffic must be processed by the Aruba MC before it can be forwarded to any other device. This allows the MC to apply stateful and application inspection policies, as well as other security features, to better secure the traffic, greatly limiting user-to-user threats in the same VLAN.

Easy scalability is provided by the Aruba MCs, where the 7280 MC supports a theoretical stateful firewall throughput of 100 Gbps.

Easier to implement security policies, because only one product is primarily used for the user traffic: Aruba Mobility Controllers. The problem with a company that has Aruba MCs for wireless and a third-party firewall is that it becomes very difficult to ensure that policies are consistent between the third-party firewall and the Aruba MCs.

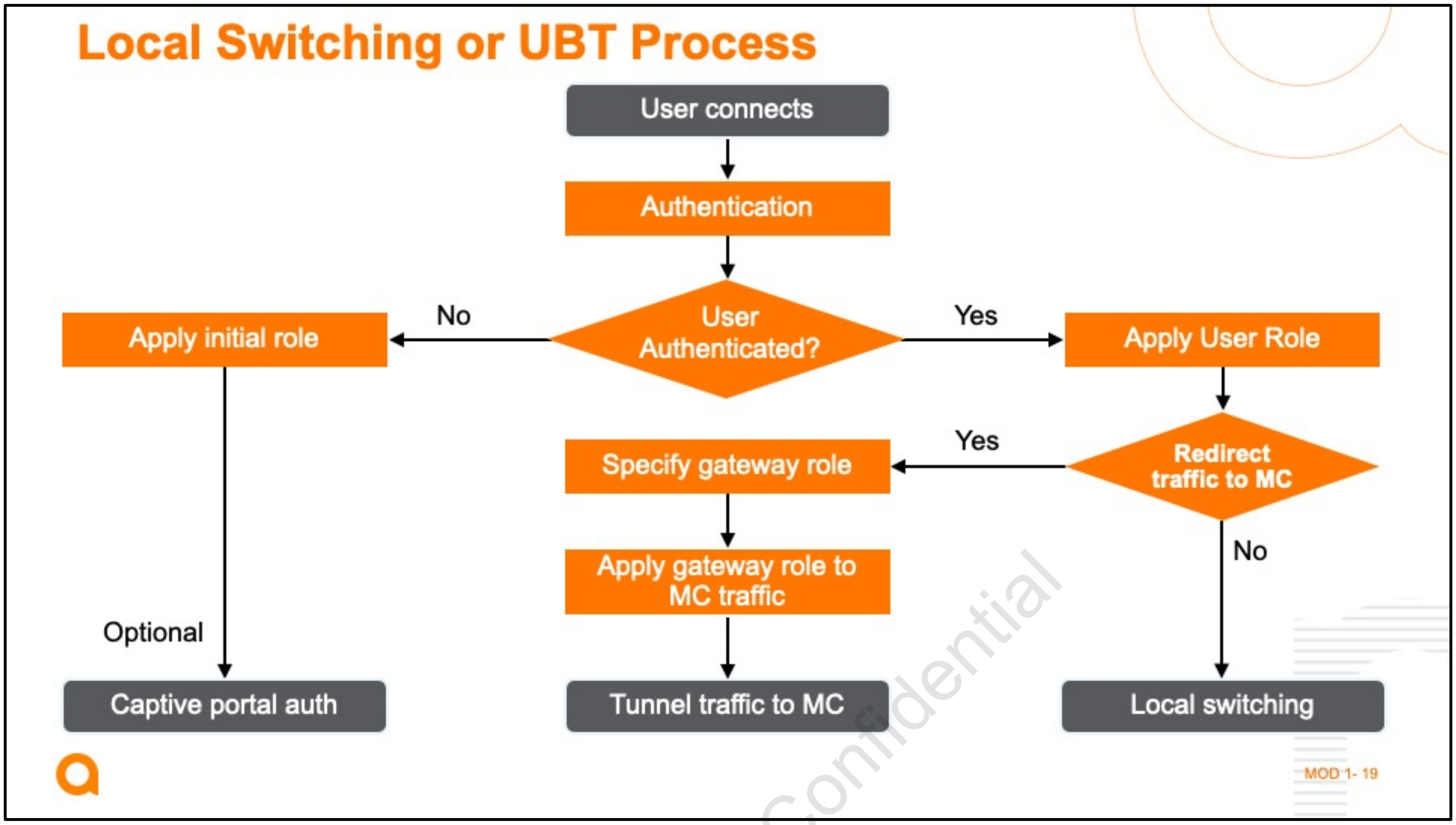


Local Switching versus UBT Summary

	Local Switching	UBT
How it works	Switch locally switches traffic	Switch tunnels traffic to Aruba MC, which processes the traffic
When to use it	• Delay-sensitive traffic: VoIP, video, etc. • Third-party firewall already installed	• Wired/wireless traffic need consistent policies • Aruba MCs are part of the existing network or new design
Advantages	• Less delay sensitive endpoint traffic	• **More secure: MC stateful firewall, application inspection, and much more** • **Traffic is controlled before a device can communicate with another device**
Disadvantages	• **Allows user-to-user connections unless filtered by ACLs (less secure)** • **Access layer switches do not support stateful firewall or application inspection features**	• Requires possible additional MCs for Ethernet bandwidth needs • Switches consume AP licenses • Requires AOS-CX or AOS switches only

MOD 1- 18

For your reference, the table summarizes the previous discussion of Local switching vs. UBT.



Local Switching or UBT Process

User connects

Authentication

User Authenticated?

No

Apply initial role

Optional

Captive portal auth

Yes

Apply User Role

Redirect traffic to MC

Yes

Specify gateway role

Apply gateway role to MC traffic

Tunnel traffic to MC

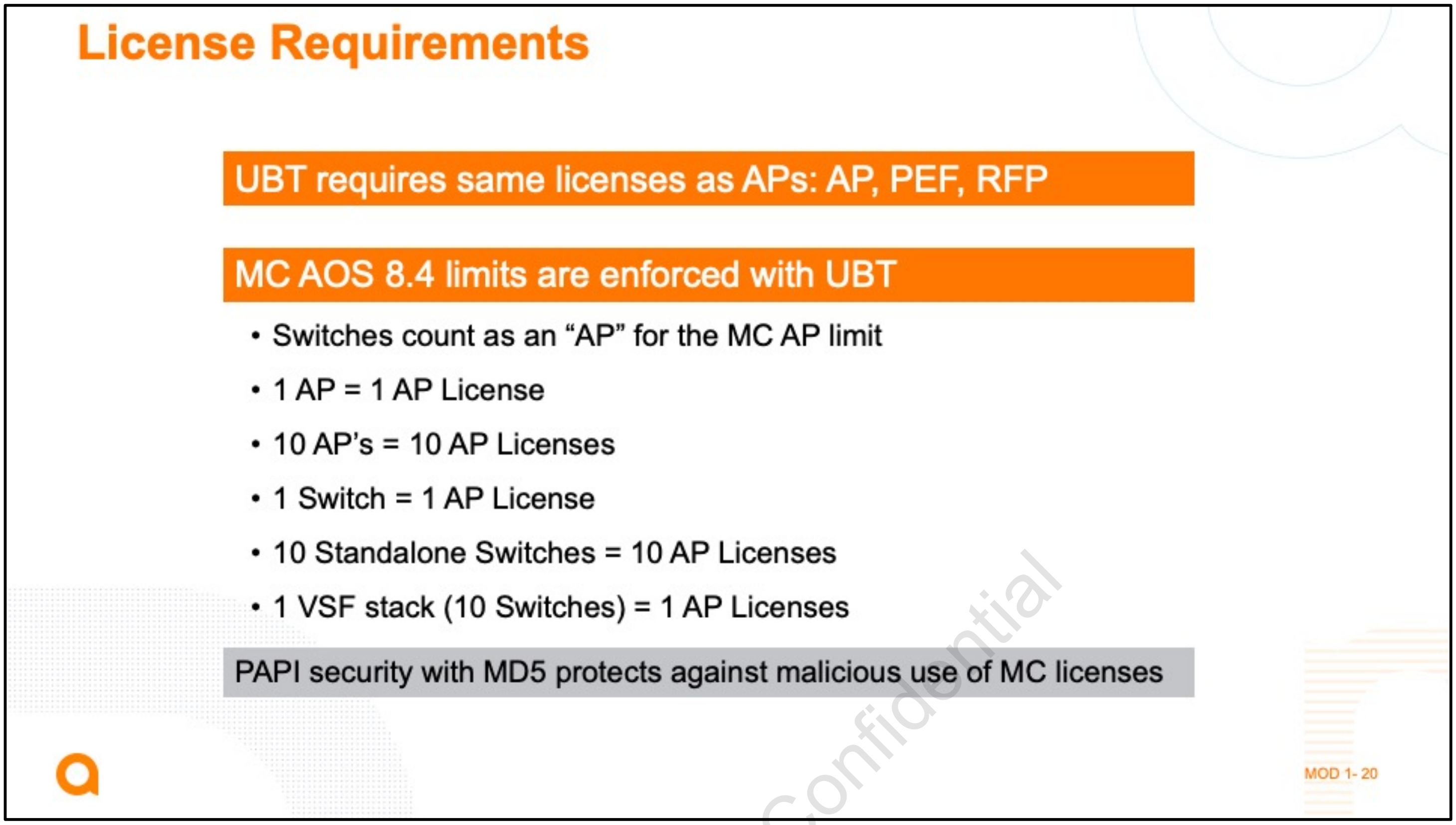
No

Local switching

MOD 1- 19

Review the figure to understand how the switch knows to switch locally or used UBT. By identifying if the redirect parameter exists, the switch can easily decide that the client traffic is tunneled to the (Mobility Controller) or is locally switched.

Important: Currently, voice traffic must use local switching.



License Requirements

UBT requires same licenses as APs: AP, PEF, RFP

MC AOS 8.4 limits are enforced with UBT

- Switches count as an “AP” for the MC AP limit
- 1 AP = 1 AP License
- 10 AP’s = 10 AP Licenses
- 1 Switch = 1 AP License
- 10 Standalone Switches = 10 AP Licenses
- 1 VSF stack (10 Switches) = 1 AP Licenses

PAPI security with MD5 protects against malicious use of MC licenses

MOD 1- 20

Initially, Aruba did not require any licensing to implement dynamic segmentation. Starting with AOS 8.4 on the MCs, the switch is not treated as an AP and will consume licenses as an AP. Note that if you set up a virtual switch (VSF) that is comprised of multiple physical AOS-CX switches, only the virtual switch will consume licenses.

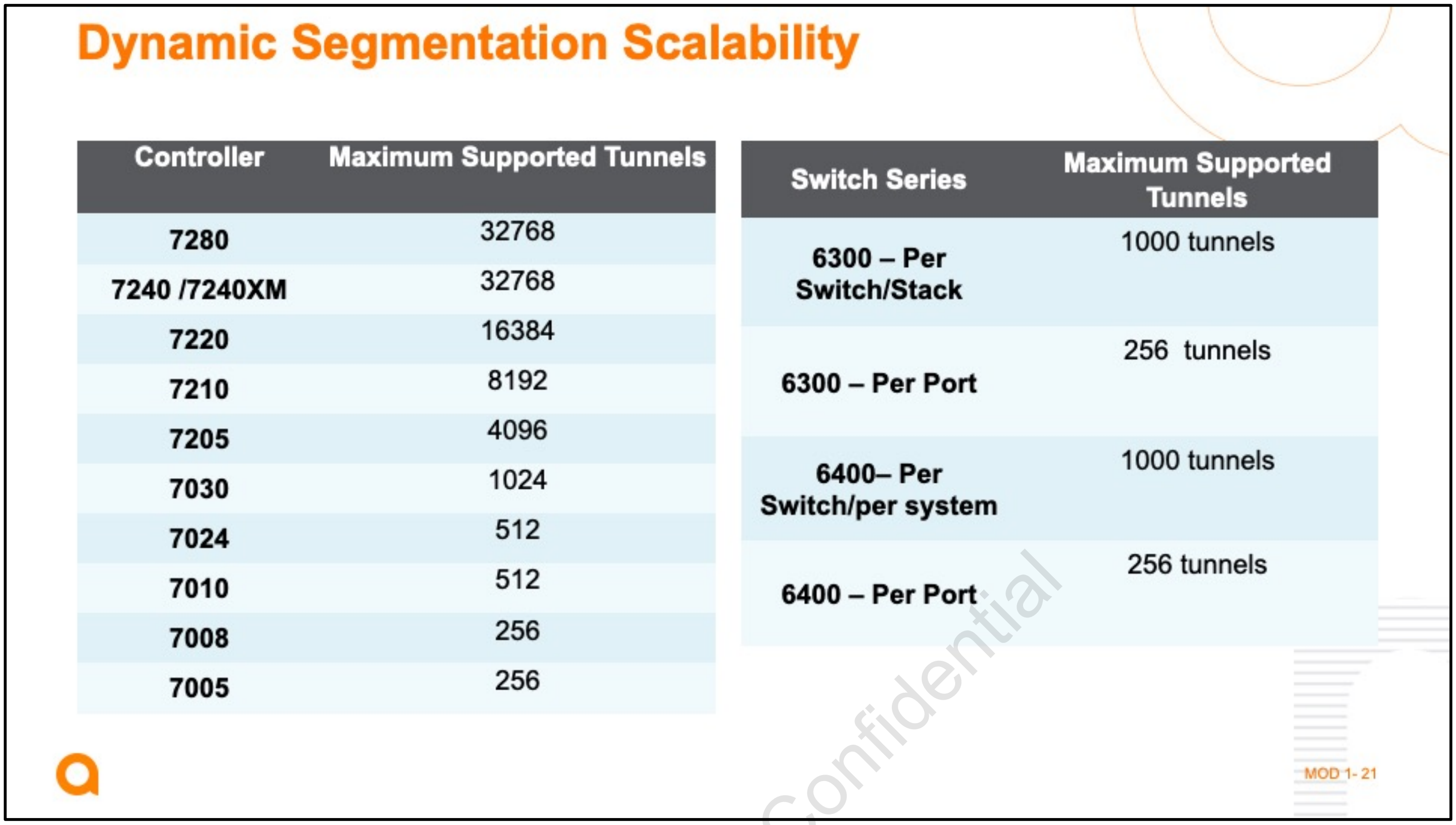
A standalone MC (or centralized MM-based solution) needs one AP license for each switch that is running dynamic segmentation. If the solution is using firewall functionality, it also needs one Policy Enforcement Firewall (PEF) license for each switch implementing dynamic segmentation.

If the MC is using the RF Protect features, it should have one RFP license for each dynamic segmentation switch as well. The RFP license is not strictly required. However, licensing works such that, when RF Protect is enabled, the MC should have the same number of AP, PEF, and RFP licenses; if it has fewer RFP licenses, it disables the excess AP and PEF licenses until it has the same count. The company might choose to deploy Aruba MCs dedicated to supporting the wired tunneled nodes. In this case, the MC would not use RF Protect, and the RFP licenses would not be required.

If the MC is part of an MM-based solution, an MM license is also required for the dynamic segmentation switch, just as it would be for an AP.

To conserve licenses, you may choose to consolidate the dynamic segmentation feature to one or just a few switches.

Important: You should always implement PAPI MD5 security to protect communications between the controller and switch, as well as protecting against malicious misuses of licenses, since each switch request consumes a license(s) on the MC.



Dynamic Segmentation Scalability

Controller	Maximum Supported Tunnels
7280	32768
7240 /7240XM	32768
7220	16384
7210	8192
7205	4096
7030	1024
7024	512
7010	512
7008	256
7005	256

Switch Series	Maximum Supported Tunnels
6300 – Per Switch/Stack	1000 tunnels
6300 – Per Port	256 tunnels
6400– Per Switch/per system	1000 tunnels
6400 – Per Port	256 tunnels

MOD 1- 21

The table shows the number of supported tunnels for both the MCs and the AOS-CX switches. Typically, the number of tunnels required to implement dynamic segmentation is not a concern if you are implementing one tunnel per port on a switch. However, if you are implementing multiple tunnels on a port, because there are multiple downstream devices (perhaps another switch or a blade server with many VMs in a data center), then you must ensure that no more than 256 tunnels are created per port.

As mentioned earlier, architects should consider the dynamic segmentation feature as they plan. In addition to the licensing requirements, they should keep in mind tunneling requirements. The MC treats each user as a separate tunnel. Architects should also remember that a single tunneled-node port could connect to multiple devices (if, for example, the port connects to an unmanaged switch), which could raise the user count. The MC must support these additional tunnels.

Finally, architects should consider the effects of the wired traffic on the firewall throughput and whether the controller can support the bandwidth; and if not, an additional controller(s) will need to be purchased.

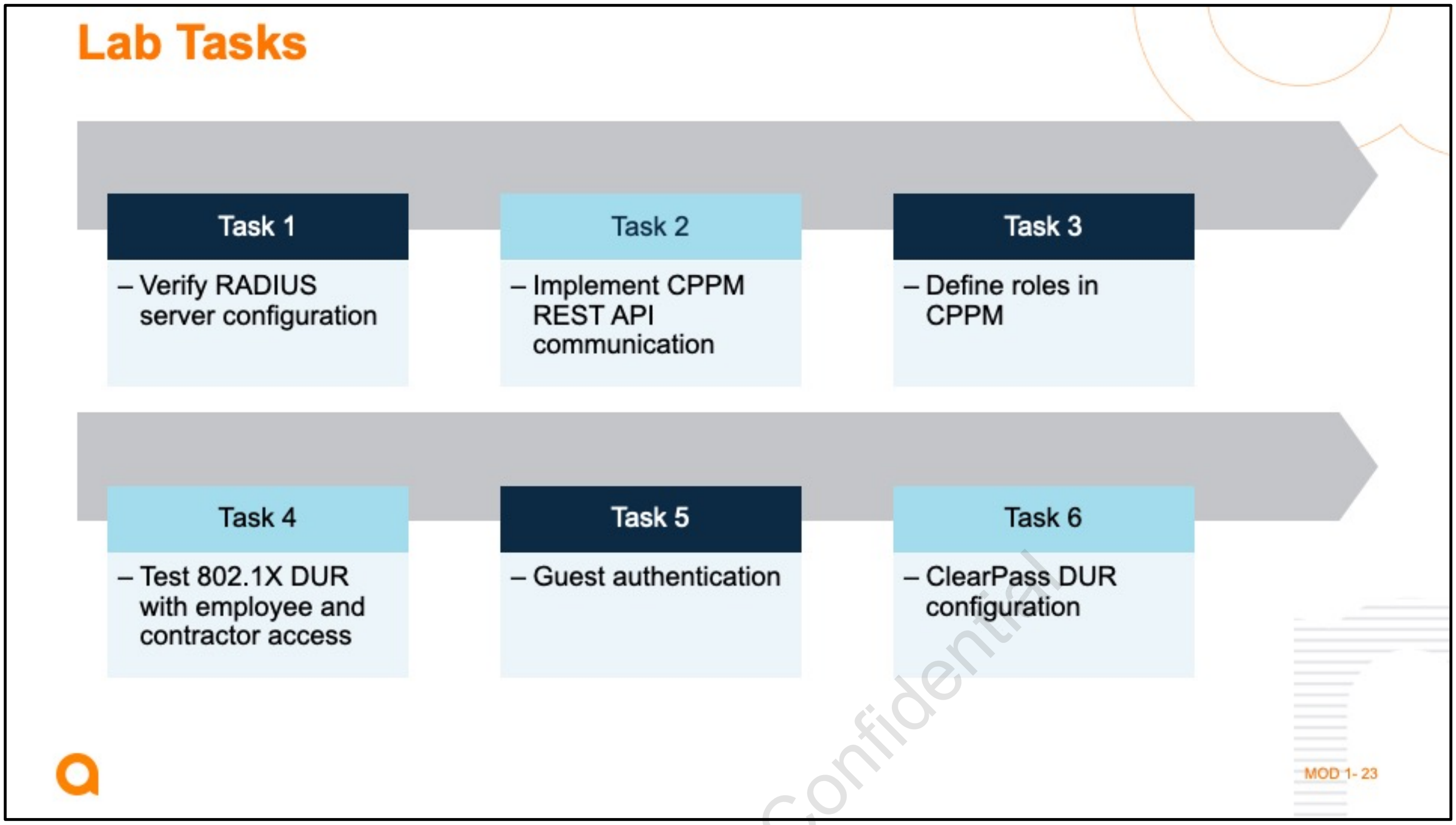
Of course, you must also implement the policies for wired traffic control. However, configuring the MC to control and forward the traffic is beyond the scope of this course and should be handled by an experienced Aruba controller administrator.

Note: Also make sure that your MCs are running up-to-date software. Minimally, ArubaOS 8.4 is required on your MCs to support user-based mode on the AOS-CX switches.

Lab Activity

Lab 6.3

It’s a good time for a lab activity.



Lab Tasks

Task 1

– Verify RADIUS server configuration

Task 2

– Implement CPPM REST API communication

Task 3

– Define roles in CPPM

Task 4

– Test 802.1X DUR with employee and contractor access

Task 5

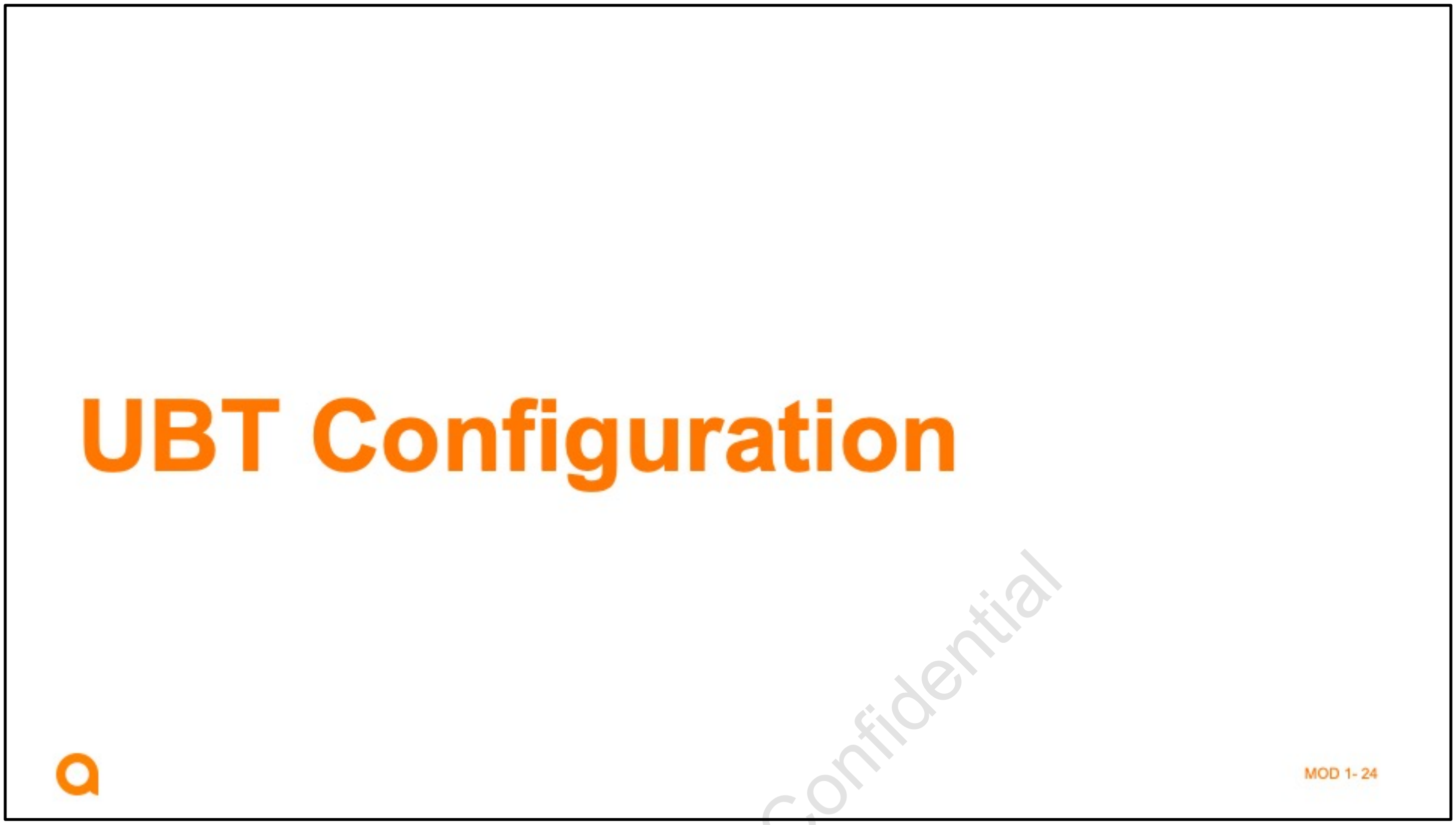
– Guest authentication

Task 6

– ClearPass DUR configuration

MOD 1- 23

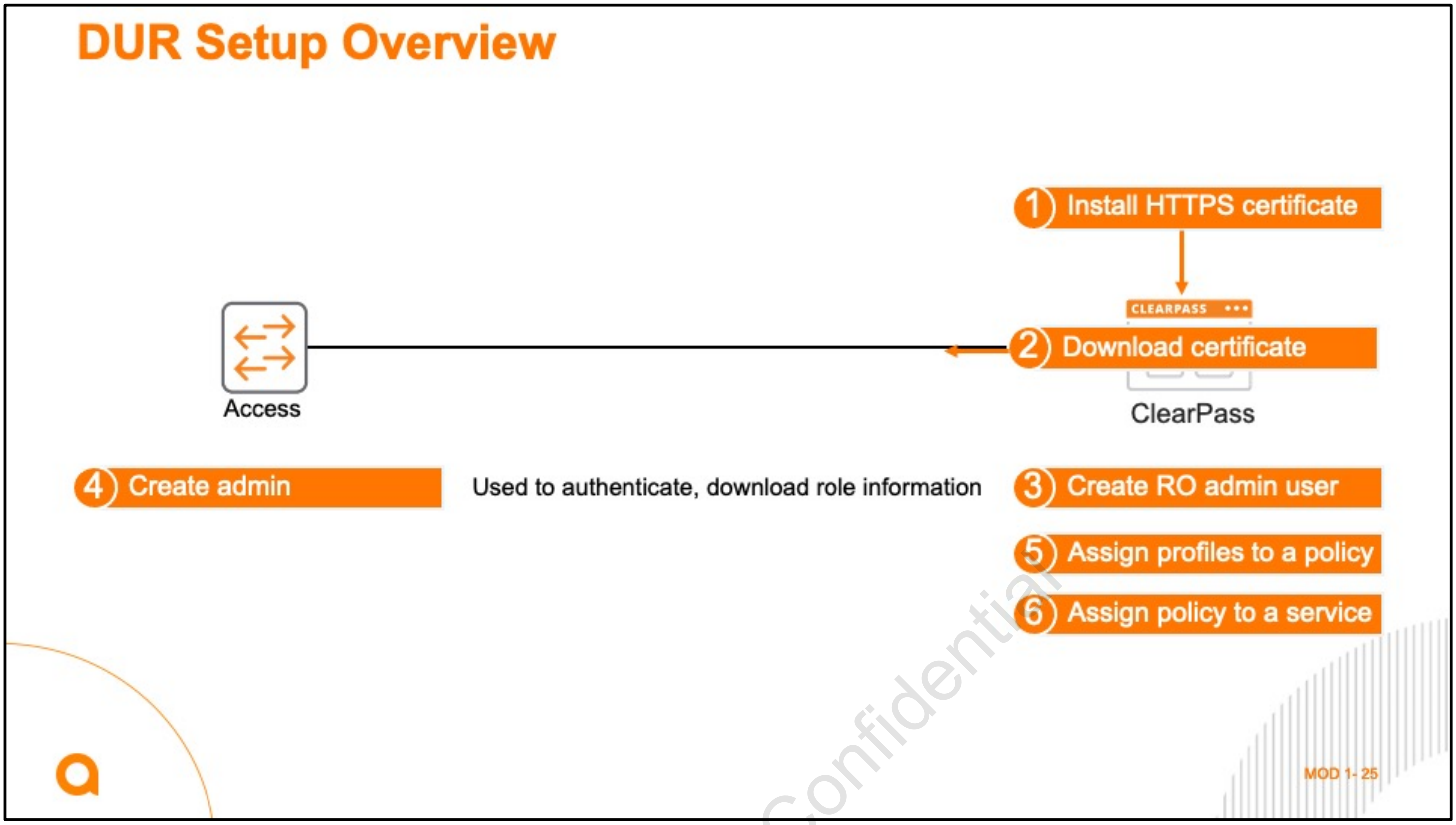
The figure provides a brief review of lab tasks. Please see your lab guide for details. When you are ready, please continue with the next section of this module.



UBT Configuration

MOD 1- 24

This section introduces you to the configuration of UBT using local user roles (LUR) and downloadable user roles (DUR)



DUR Setup Overview

1 Install HTTPS certificate

2 Download certificate

3 Create RO admin user

4 Create admin

5 Assign profiles to a policy

6 Assign policy to a service

Used to authenticate, download role information

MOD 1- 25

Recall that you define DURs on the ClearPass server, which are downloaded to the switch after successful authentication. Because this happens via a REST API using an SSL connection, you must install an HTTPS certificate on the edge switch. Plus, a username/password must be configured to authenticate the switch with ClearPass. Here's how it works:

First, ensure that ClearPass has obtained a valid HTTPS certificate

Automatically download the certificate from ClearPass (Requires ClearPass v6.7.8+ or a later)

On ClearPass, create a read-only admin user. The switch uses this to authenticate and download role information

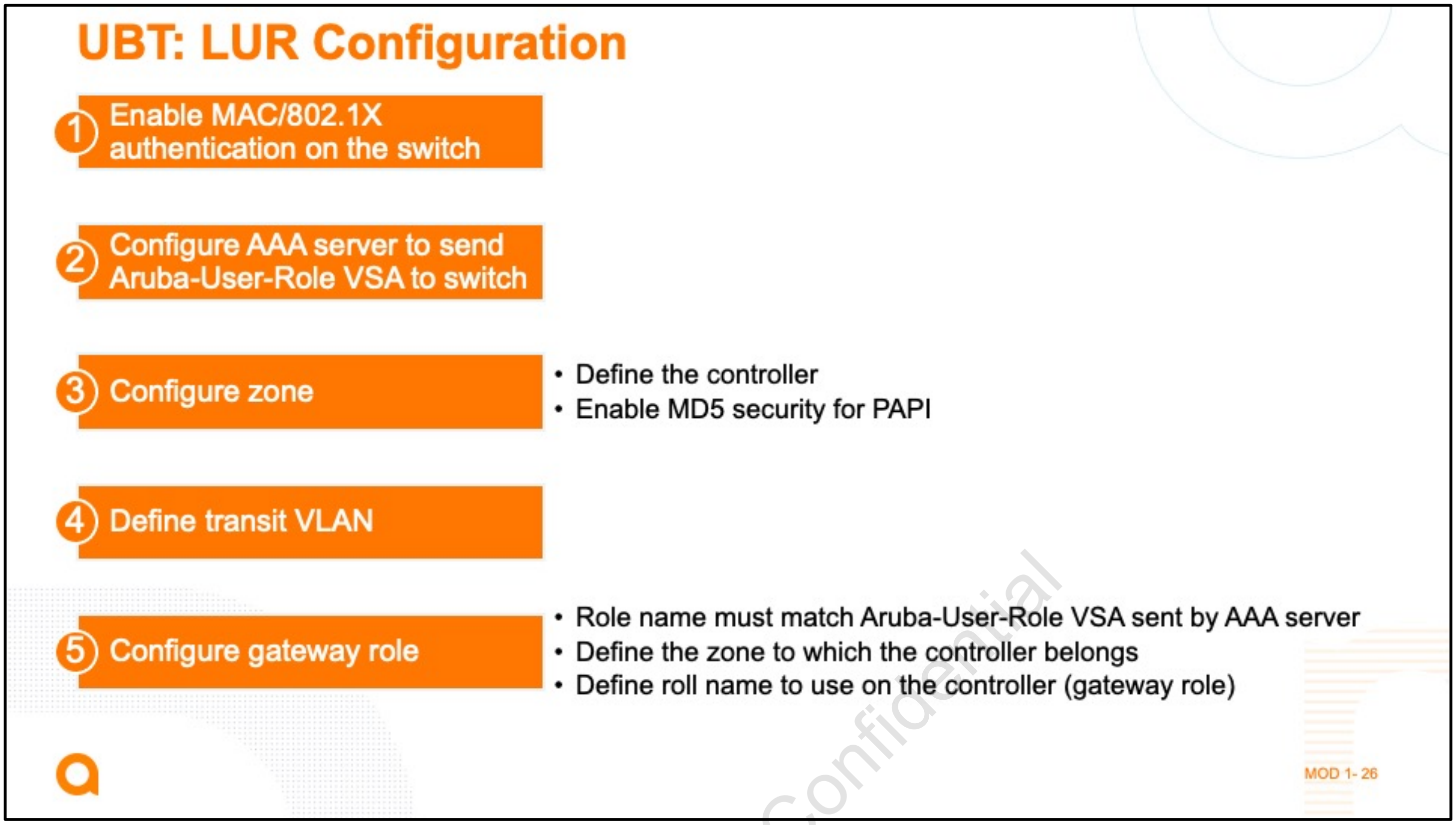
On the switch, create a user that matches the one created on ClearPass. Enable downloadable user roles and configure the downloadable user roles profiles

Assign the profiles to a policy and the policy to a service on the ClearPass server

You can do this with a cluster of controllers, or with a stand-alone controller.

Note: Because the login process from the AOS-CX switch uses HTTPS, the switch must validate the ClearPass server certificate. On the older AOS switches, they would automatically download and install the ClearPass certificate. Unfortunately, in AOS-CX 10.4, this process is not automatic. You must access the ClearPass server, download its SSL certificate, and install it on the AOS-CX switch.

Important: UBT/local switching and user roles are two different, but related concepts. They are both used to implement dynamic segmentation. With Aruba AOS-CX switches, some devices could be locally switched with a pre-define local role, perhaps for VoIP phones; but for other devices, their role information could be downloaded from ClearPass and the traffic tunneled to an Aruba MC. It’s flexible in how you want to set it up. However, Aruba recommends that if you are using ClearPass and have either AOS-CX or the older AOS switches, that you implement DUR because of the scalability and ease of centralized policy management that it provides.



UBT: LUR Configuration

1. Enable MAC/802.1X authentication on the switch
2. Configure AAA server to send Aruba-User-Role VSA to switch
3. Configure zone
   - Define the controller
   - Enable MD5 security for PAPI
4. Define transit VLAN
5. Configure gateway role
   - Role name must match Aruba-User-Role VSA sent by AAA server
   - Define the zone to which the controller belongs
   - Define roll name to use on the controller (gateway role)

MOD 1- 26

To configure UBT with Local User Roles you must first configure 802.1X, MAC-Auth, and/or captive portal on your AOS-CX switch, along with connectivity to an AAA server, like ClearPass. You learned about Steps 1 and 2 in earlier modules –this configuration is not shown here, so we begin with step 3.

3. Configure zone

To set up UBT LUR, you first need to define a zone:

```
switch(config)# ubt zone <zone-name> vrf <VRF-name>
switch(config-ubt-zone)# primary-controller ip <IP-address>
switch(config-ubt-zone)# backup-controller ip <IP-address>
switch(config-ubt-zone)# papi-security-key <key>
switch(config-ubt-zone)# enable
switch(config)# ip source-interface ubt {interface <IFNAME> | <IPV4-ADDR>} [vrf <VRF-NAME>]
```

First you create the zone and define the VRF instance that will be used. Next, you define the primary, and optionally, backup IP address of the controller(s). Unlike the APs, the switches currently do not support Control Plane Security (CPSec). Instead, you can optionally protect the PAPI control messages using MD5 signatures. They key that you specify must match what the MC(s) have configured. Last, you need to enable the zone. Optionally you can define which

IP address the switch should use when tunneling the GRE packets (by default this is based on the local routing table).

Note: In AOS-CX 10.4, currently only one zone is supported per switch.

To verify your configuration, there are various show commands that you can use:

```
switch# show ubt
Zone Name : my-zone
Primary Controller : 10.116.51.10
Backup Controller : 10.116.51.11
SAC HeartBeat Interval : 1
UAC KeepAlive Interval : 60
Reserved VLAN Identifier : 4094
VRF Name : my-vrf
Admin State : ENABLED
PAPI Security Key :
AQBapdxySvGPvdTlkYn1/naKX4O3jKHrm28xLYfO6mLOK499BwAAAHdJp/b
L4FE=
```

```
switch# show ubt information
SAC Information
Active : 10.116.51.12
Standby : 10.116.51.13
Node List Information
Cluster Name : my-cluster
Node List
10.1.1.1
10.1.1.2
10.1.1.3
Bucket Map Information
Bucket Map Active : [0...255]
Bucket ID A-UAC     S-UAC     Connectivity
```

```
--------------------------------------------------------
0        10.1.1.1  10.1.1.2  L2
1        10.1.1.2  10.1.1.3  L2
2        10.1.1.3  10.1.1.1  L2
```

4. Define Transit VLAN

To define the transit VLAN, use the following configuration:

```
Switch(config)# vlan <VLAN-ID>
Switch(config-vlan)# exit
Switch(config)# ubt-client-vlan <VLAN-ID?
```

Remember that the transit VLAN is what the switch will use when tunneling the user traffic to an MC (if the switch is not performing switching of the traffic).

Note: For AOS-CX 104, Aruba, only supports one zone.

5. Configuring a gateway role

Next, you need to define the role names locally on the switch:

```
switch(config)# port-access role <role-name>
switch(config-pa-role)# gateway-zone zone <zone-name> gateway-role
<controller-role-name>
```

The role name in the port-access role command must exactly match the Aruba-User-Role VSA that the AAA server, like ClearPass, will pass back to the switch. In the role configuration, you must define the zone that this is used for (currently only one zone is supported) and then the role name that is passed to the MC. Note that the switch role name and the MC role name could be the same, or different, names. The controller role name is also case-sensitive. See the Security Guide for other configuration options for the role. For example, you could define a policy to define a pre-authentication VLAN, ACL, QoS and other parameters that should be used before the user is authenticated.

Enable user authentication

After you’ve configured the role information, you then need to set up user authentication, like MAC or 802.1X (captive portal is also supported). Here's an example with 802.1X:

```
Switch(config)# aaa authentication port-access dot1x authenticator auth-method eap-radius
Switch(config)# aaa authentication port-access dot1x authenticator radius server-group radius
Switch(config)# aaa authentication port-access dot1x authenticator enable
Switch(config)# interface <interface-id>
Switch(config-if)# aaa authentication port-access dot1x authenticator
Switch(config-if-dot1x-auth)# enable
Switch(config-if-dot1x-auth)# exit
```

Optionally, you can define a pre-authentication role on the interface—this is applied to the user's traffic prior to authentication:

```
Switch(config-if)# aaa authentication port-access preauth-role <role-name>
```

Also, if you need to exempt LLDP, CDP, and/or BPDU traffic, use the following two commands:

```
switch(config-if)# aaa authentication port-access allow-lldp-bpdu
switch(config-if)# aaa authentication port-access allow-cdp-bpdu
```

This configuration might be necessary for devices like a phone or camera where LLDP (or Cisco's CDP) must be exchanged for the device to operate correctly prior to authentication occurring on the port.

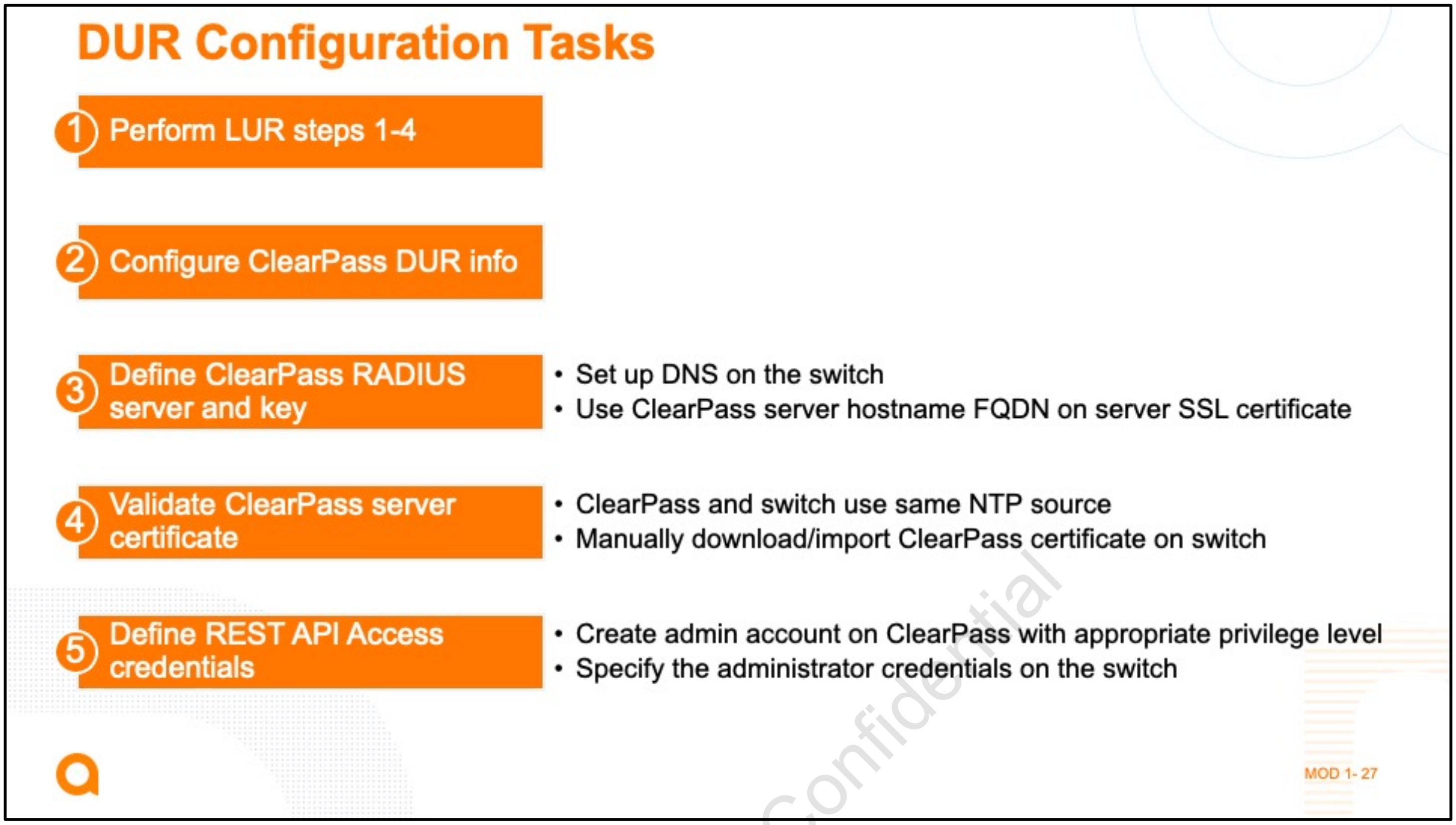
Configure your AAA server role policy

Last, you need to define a policy on your AAA server to return the respective role name for the user. This is done using an Aruba VSA called Aruba-User-Role. Remember that the role names are case sensitive!

You can also see a list of the users and the roles assigned to them when verifying the operation of LUR:

```
switch# show ubt users all
Downloaded user roles are preceded by *
Port   Mac Address       Tunnel Status  Secondary UserRole  Failure Reason
```

```
------------------------------------------------------------------------
1/1/1  00:00:00:11:12:03  activated      authenticated       ---/---
```



DUR Configuration Tasks

1. Perform LUR steps 1-4
2. Configure ClearPass DUR info
3. Define ClearPass RADIUS server and key
   - Set up DNS on the switch
   - Use ClearPass server hostname FQDN on server SSL certificate
4. Validate ClearPass server certificate
   - ClearPass and switch use same NTP source
   - Manually download/import ClearPass certificate on switch
5. Define REST API Access credentials
   - Create admin account on ClearPass with appropriate privilege level
   - Specify the administrator credentials on the switch

MOD 1- 27

Steps 1-4 of the LUR configuration task are performed here as well (shown here):

Enable MAC/802.1X authentication on the switch

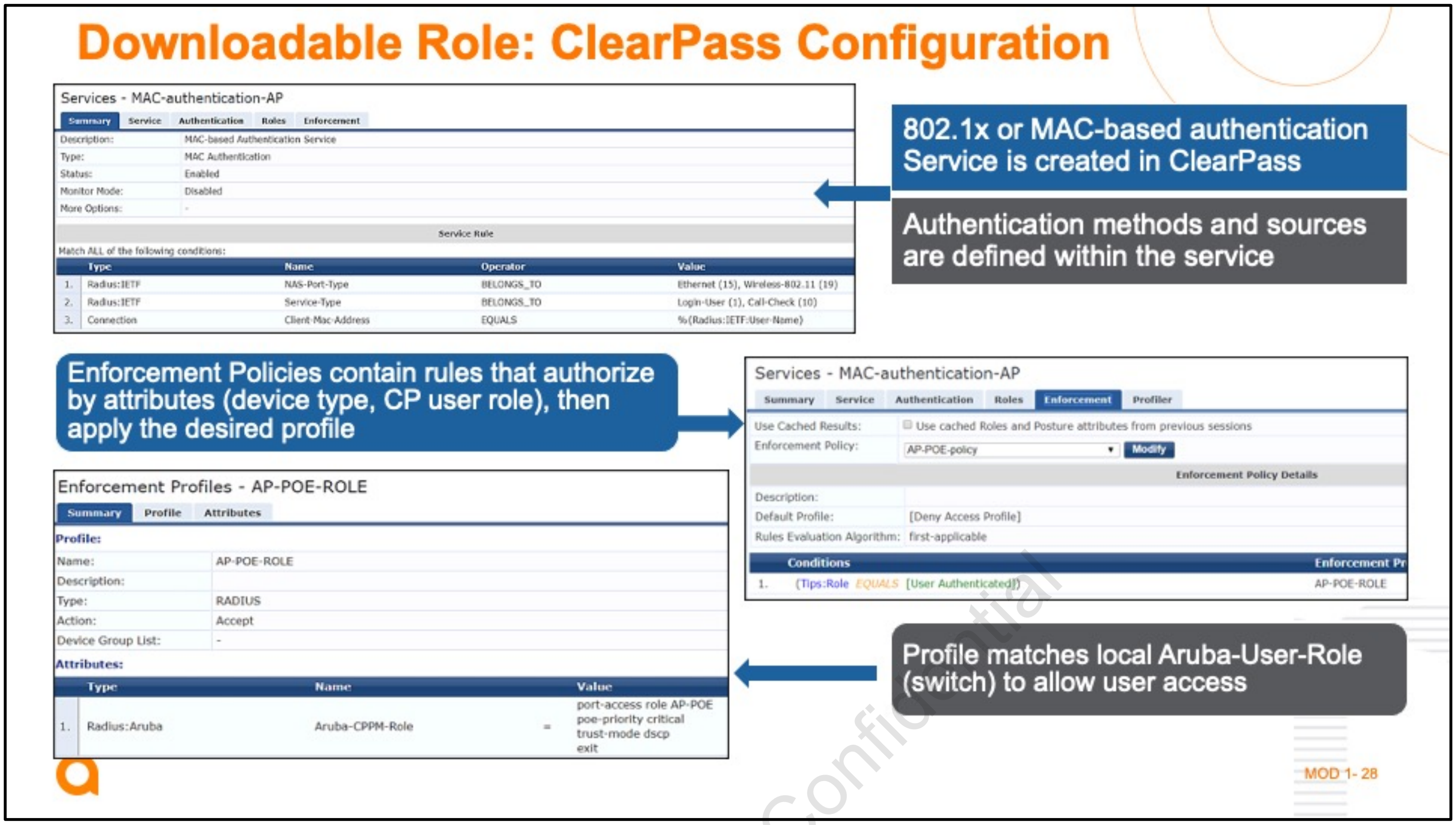
Configure the ClearPass server to send Aruba-User-Role (VSA) to switch

Configure the zone on the switch:

- Define the controller
- Enable MD5 security for PAPI

Define the transit VLAN on the switch and the Mobility Controller

Once this is done, you need to perform the steps 2-5 listed in the above slide. These steps are covered in the next pages.



Downloadable Role: ClearPass Configuration

Services - MAC-authentication-AP

Description:	MAC-based Authentication Service
Type:	MAC Authentication
Status:	Enabled
Monitor Mode:	Disabled
More Options:	-

Service Rule

Match ALL of the following conditions:

	Type	Name	Operator	Value
1.	Radius:IETF	NAS-Port-Type	BELONGS_TO	Ethernet (15), Wireless-802.11 (19)
2.	Radius:IETF	Service-Type	BELONGS_TO	Login-User (1), Call-Check (10)
3.	Connection	Client-Mac-Address	EQUALS	%{Radius:IETF:User-Name}

802.1x or MAC-based authentication Service is created in ClearPass

Authentication methods and sources are defined within the service

Enforcement Policies contain rules that authorize by attributes (device type, CP user role), then apply the desired profile

Services - MAC-authentication-AP

Use Cached Results:	□ Use cached Roles and Posture attributes from previous sessions
Enforcement Policy:	AP-POE-policy ▾ Modify

Enforcement Policy Details

Description:	
Default Profile:	[Deny Access Profile]
Rules Evaluation Algorithm:	first-applicable

	Conditions	Enforcement Pr
1.	(Tips:Role EQUALS [User Authenticated])	AP-POE-ROLE

Enforcement Profiles - AP-POE-ROLE

Profile:

Name:	AP-POE-ROLE
Description:	
Type:	RADIUS
Action:	Accept
Device Group List:	-

Attributes:

	Type	Name		Value
1.	Radius:Aruba	Aruba-CPPM-Role	=	port-access role AP-POE poe-priority critical trust-mode dscp exit

Profile matches local Aruba-User-Role (switch) to allow user access

MOD 1- 28

The process for DUR is very much the same as LUR. Instead of having the user role configured on the edge switch, however, the user role is configured on ClearPass and sent to the edge switch after successful authentication. The method for sending the user role is through REST API using an SSL connection. This means that a HTTPS certificate has to be installed on the edge switch. In addition, a username/password has to be configured to authenticate the switch with ClearPass.

Let’s have a look how this is done:

First, ensure that ClearPass has obtained a valid HTTPS certificate =

Automatically download the certificate from ClearPass (ClearPass 6.7.8+ or a later version is required)

On ClearPass, create a read only admin user that will be used by the switch for authentication in order to download the actual role information

On the switch, create the user entry for the read only user and enable downloadable user roles and configure the downloadable user roles profiles

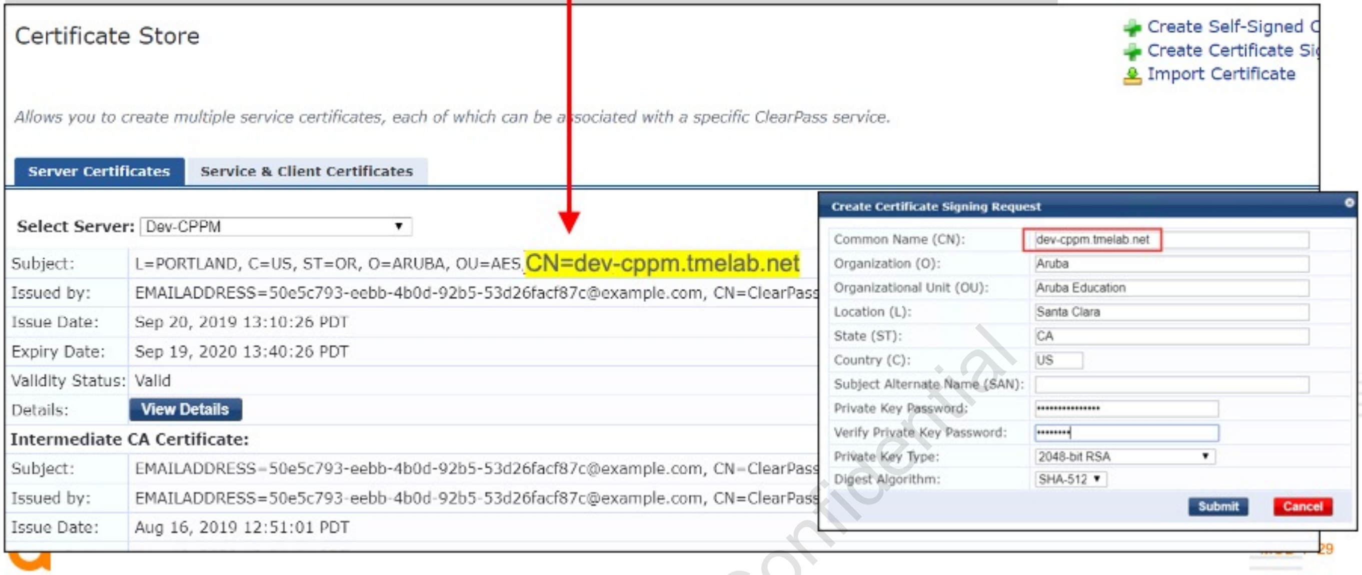
Assign the profiles to a policy and the policy to a service on the ClearPass server

Note: LUR should be used if you have a generic AAA server product. Either LUR or DUR can be used if you have ClearPass.

This slide lists the authentication service/enforcement policy and profile configuration on Clearpass. Please see the ClearPass documentation for more information on configure services, enforcement policies, and enforcement profiles.

Downloadable Role: ClearPass HTTPS Server Cert

```
ip dns server-address <IP-address> vrf <VRF-name>
radius-server host dev-cppm.tmelab.net key <key> vrf <VRF-name>
```



Certificate Store

Allows you to create multiple service certificates, each of which can be associated with a specific ClearPass service.

Select Server: Dev-CPPM

Subject:	L=PORTLAND, C=US, ST=OR, O=ARUBA, OU=AES,CN=dev-cppm.tmelab.net
Issued by:	EMAILADDRESS=50e5c793-eebb-4b0d-92b5-53d26facf87c@example.com, CN=ClearPass
Issue Date:	Sep 20, 2019 13:10:26 PDT
Expiry Date:	Sep 19, 2020 13:40:26 PDT
Validity Status:	Valid
Details:	View Details

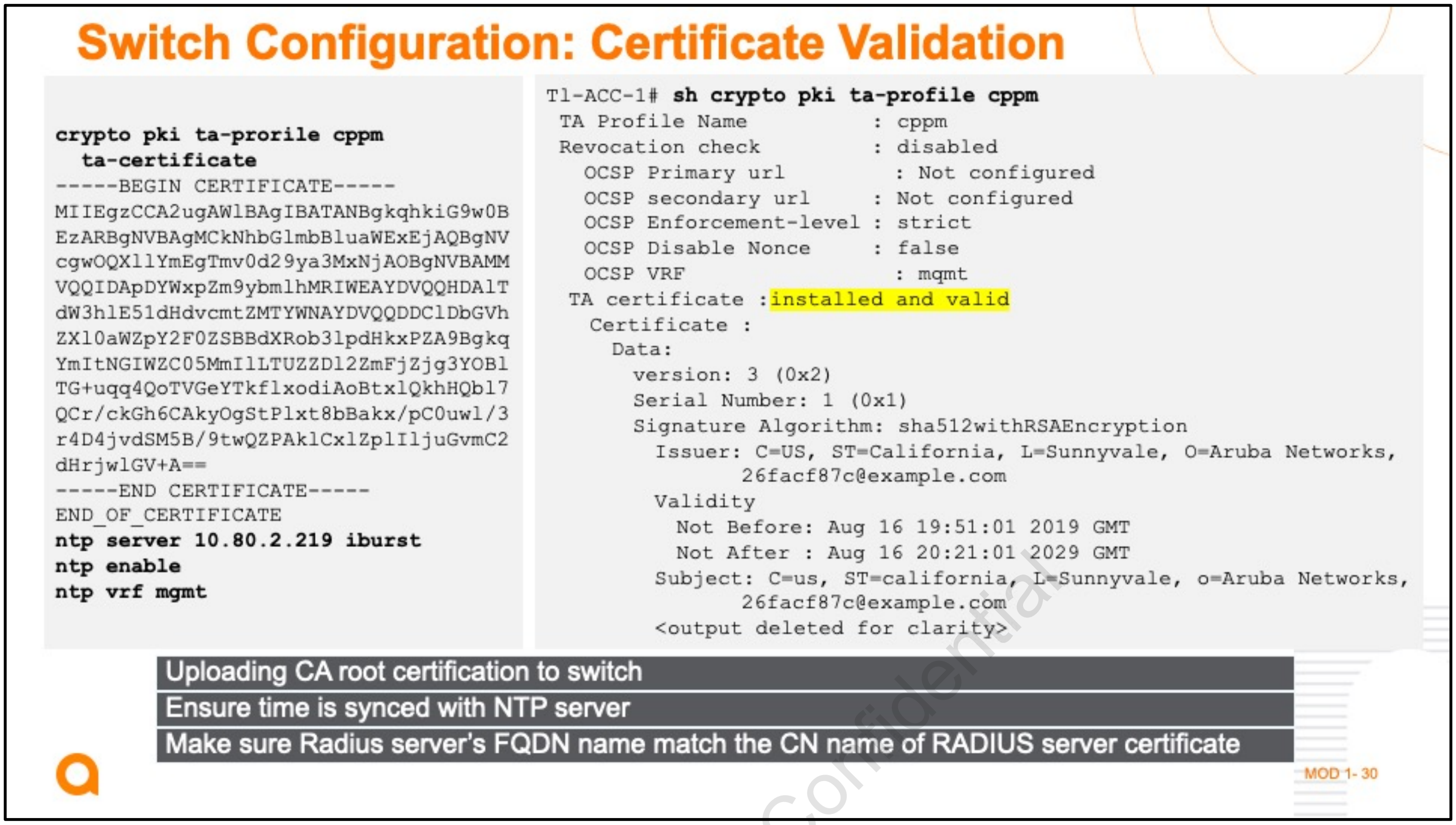
Intermediate CA Certificate:

Subject:	EMAILADDRESS=50e5c793-eebb-4b0d-92b5-53d26facf87c@example.com, CN=ClearPass
Issued by:	EMAILADDRESS=50e5c793-eebb-4b0d-92b5-53d26facf87c@example.com, CN=ClearPass
Issue Date:	Aug 16, 2019 12:51:01 PDT

Create Certificate Signing Request

Common Name (CN):	dev-cppm.tmelab.net
Organization (O):	Aruba
Organizational Unit (OU):	Aruba Education
Location (L):	Santa Clara
State (ST):	CA
Country (C):	US
Subject Alternate Name (SAN):	
Private Key Password:	••••••••••••••••
Verify Private Key Password:	••••••••
Private Key Type:	2048-bit RSA
Digest Algorithm:	SHA-512

Make sure radius-server host name on configured on the switch matches the Common Name (CN) field of ClearPass HTTPS server certificate. If a FQDN is used, you will need to specify a DNS server that the switch should use to resolve the server's name to an IP address.



Switch Configuration: Certificate Validation

```
crypto pki ta-prorile cppm
  ta-certificate
-----BEGIN CERTIFICATE-----
MIIEgzCCA2ugAWlBAgIBATANBgkqhkiG9w0B
EzARBgNVBAgMCkNhbGlmbBluaWExEjAQBgNV
cgwOQXllYmEgTmv0d29ya3MxNjAOBgNVBAMM
VQQIDApDYWxpZm9ybmlhMRIWEAYDVQQHDAlT
dW3hlE51dHdvcmtZMTYWNAYDVQQDDClDbGVh
ZXl0aWZpY2F0ZSBBdXRob3lpdHkxPZA9Bgkq
YmItNGIWZC05MmIlLTUZZDl2ZmFjZjg3YOBl
TG+uqq4QoTVGeYTkflxodiAoBtxlQkhHQbl7
QCr/ckGh6CAkyOgStPlxt8bBakx/pC0uwl/3
r4D4jvdSM5B/9twQZPAklCxlZplIljuGvmC2
dHrjwlGV+A==
-----END CERTIFICATE-----
END_OF_CERTIFICATE
ntp server 10.80.2.219 iburst
ntp enable
ntp vrf mgmt
```

```
T1-ACC-1# sh crypto pki ta-profile cppm
 TA Profile Name          : cppm
 Revocation check         : disabled
   OCSP Primary url         : Not configured
   OCSP secondary url     : Not configured
   OCSP Enforcement-level : strict
   OCSP Disable Nonce     : false
   OCSP VRF                 : mgmt
  TA certificate :installed and valid
   Certificate :
     Data:
       version: 3 (0x2)
       Serial Number: 1 (0x1)
       Signature Algorithm: sha512withRSAEncryption
         Issuer: C=US, ST=California, L=Sunnyvale, O=Aruba Networks,
                26facf87c@example.com
         Validity
           Not Before: Aug 16 19:51:01 2019 GMT
           Not After : Aug 16 20:21:01 2029 GMT
         Subject: C=us, ST=california, L=Sunnyvale, o=Aruba Networks,
                26facf87c@example.com
         <output deleted for clarity>
```

Uploading CA root certification to switch

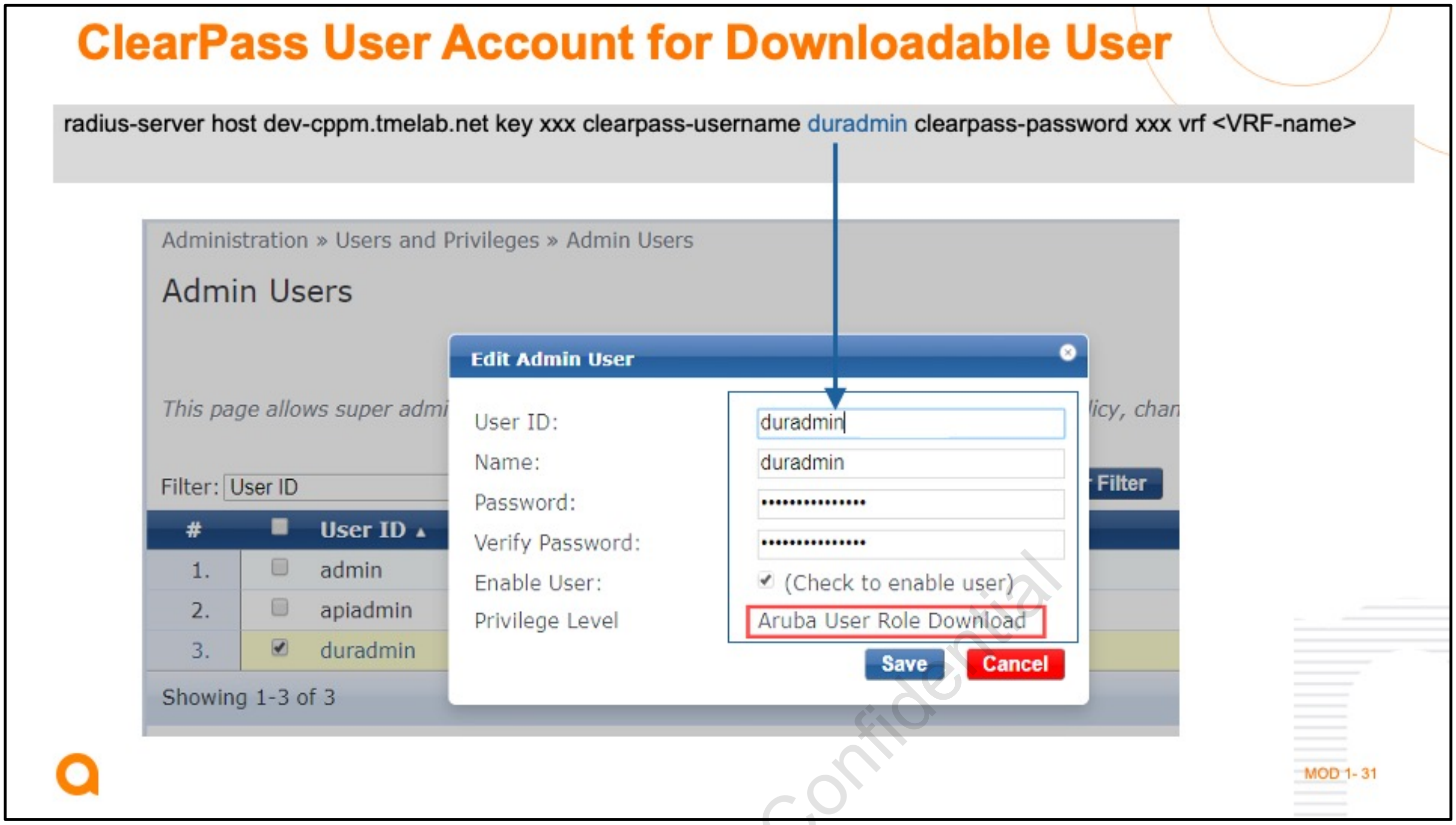
Ensure time is synced with NTP server

Make sure Radius server's FQDN name match the CN name of RADIUS server certificate

MOD 1- 30

ClearPass will send a user role to the edge switch after successful authentication. The method for sending the user role is through the REST API using an SSL connection. This means that a valid HTTPS certificate has to be installed on the edge switch: one from a trusted Certificate Authority (CA) of both the switch and the ClearPass server.

You will also need to configure NTP to ensure the certificate for switch is valid based on the current date and time and the time range defined on the certificate. If the switch's time falls outside the range of the time on the certificate, the certificate is automatically invalidated.



ClearPass User Account for Downloadable User

```
radius-server host dev-cppm.tmelab.net key xxx clearpass-username duradmin clearpass-password xxx vrf <VRF-name>
```

Administration » Users and Privileges » Admin Users

Admin Users

This page allows super admi ... licy, chan

Filter: User ID

#	□	User ID ▲
1.	□	admin
2.	□	apiadmin
3.	☑	duradmin

Showing 1-3 of 3

Edit Admin User

User ID: duradmin

Name: duradmin

Password: •••••••••••••••

Verify Password: •••••••••••••••

Enable User: ☑ (Check to enable user)

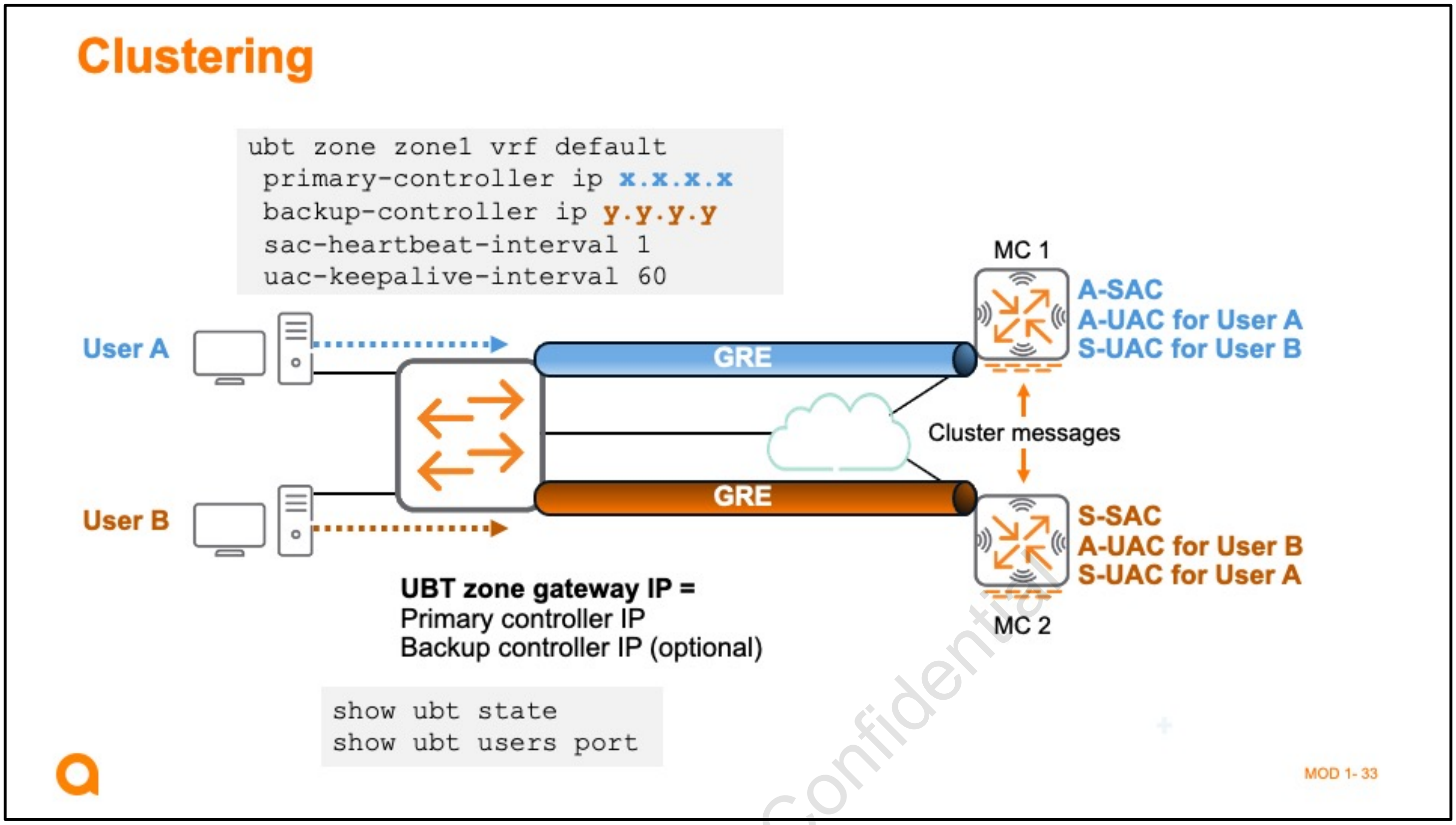
Privilege Level: Aruba User Role Download

In addition, a username and password have to be configured to authenticate the switch with the ClearPass server. This is defined on the ClearPass server, but referenced on the switch. The privilege level can be “Aruba User Role Download” or “Read-Only Administrator”.

High Availability

MOD 1- 32

Let’s learn about High Availability.



Clustering

```
ubt zone zone1 vrf default
 primary-controller ip x.x.x.x
 backup-controller ip y.y.y.y
 sac-heartbeat-interval 1
 uac-keepalive-interval 60
```

UBT zone gateway IP =
Primary controller IP
Backup controller IP (optional)

```
show ubt state
show ubt users port
```

MOD 1- 33

When the switch uses role-based tunneled-node, it can tunnel traffic to clustered MCs. Typically the switches will be connecting to a cluster of MCs. Clustering enables load balancing and failover capabilities over other redundancy solutions.

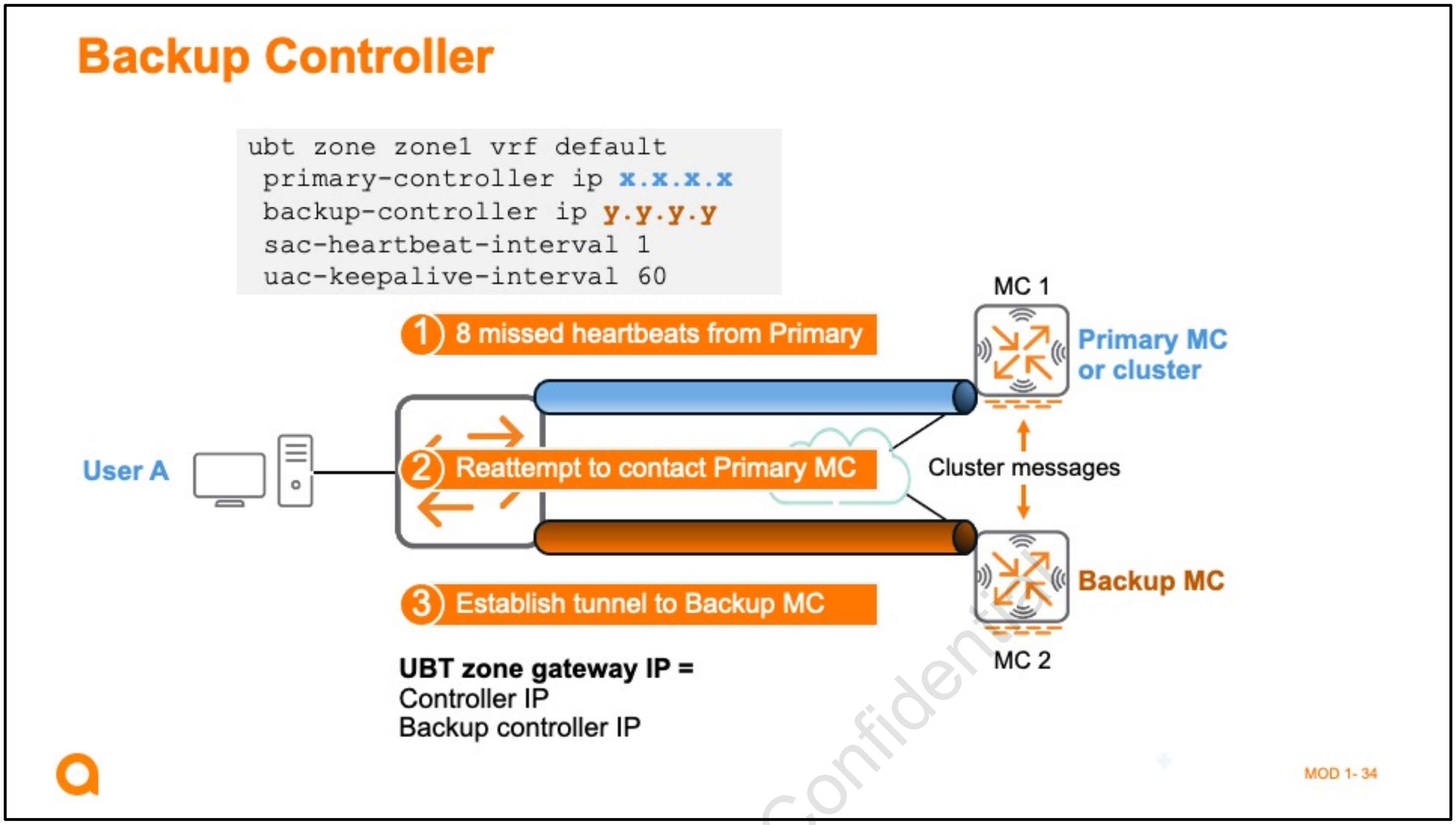
With this method, again, you only need to set a single tunnel node server (controller) IP address on the AOS-CX switch. At least define one IP address of a cluster member, and optionally a secondary address as a backup. When the switch contacts a member of the cluster, it receives assignments with the actual IP addresses of two MCs. One is the switch's active Switch Anchor Controller (A-SAC), and the other is the switch's standby SAC (S-SAC). The switch establishes PAPI tunnels with both controllers.

In addition, the switch receives a bucket-map. When the switch starts tunneling a new device's traffic, the switch hashes information about the session. It looks up the hash in the bucket map to determine the correct active User Anchor Controller (A-UAC) and standby UAC (S-UAC) for that device. The switch tunnels the device's traffic to the A-UAC, but the state for the session is synchronized between the A-UAC and the S-UAC, enabling more seamless failover. (The degree of synchronized information depends on the device's needs.) In addition, because the cluster dictates where the switch tunnels each user's traffic, load balancing is improved.

If the switch's A-SAC or any of its A-UACs fail, the cluster can detect that failure very quickly and signal to the switch to failover to the appropriate standby tunnel, rather than the switch having to wait for its heartbeats to time out.

In the above example, notice that user A and B's active and standby tunnels are not on the same controller. User A's active tunnel is on the left controller while User B's active tunnel is on the right controller.

As you have seen, the AOS-CX switch can participate in high availability solutions with a single controller IP address specified in its tunneled node server profile. For additional redundancy, the primary and secondary addresses can be controllers in two different clusters. The zone profile can also include a backup controller, though, specified with the configuration shown in the figure



Backup Controller

```
ubt zone zone1 vrf default
 primary-controller ip x.x.x.x
 backup-controller ip y.y.y.y
 sac-heartbeat-interval 1
 uac-keepalive-interval 60
```

1) 8 missed heartbeats from Primary

2) Reattempt to contact Primary MC

3) Establish tunnel to Backup MC

UBT zone gateway IP =
Controller IP
Backup controller IP

As you have seen, the AOS-CX switch can participate in high availability solutions with a single controller IP address (not part of a cluster) specified in its tunneled node server profile. This profile can also include a backup controller, though, as shown in the figure.

These commands were discussed previously. The backup controller allows the switch to failover to an MC (or cluster) in a different subnet from its primary MC (or cluster).

As discussed earlier, the switch and its MC—or any of its SACs and UACs when clustering is used—send each other keepalives over the GRE tunnel every second to assess the tunnel's health. If the switch fails to hear keepalives over the keepalive timeout period, it attempts to re-establish communications with the primary controller. In this scenario, though, the company has a standalone MC at a different location. After failing to connect to the primary MC (or cluster), the switch establishes a tunnel with the backup controller at the different location.

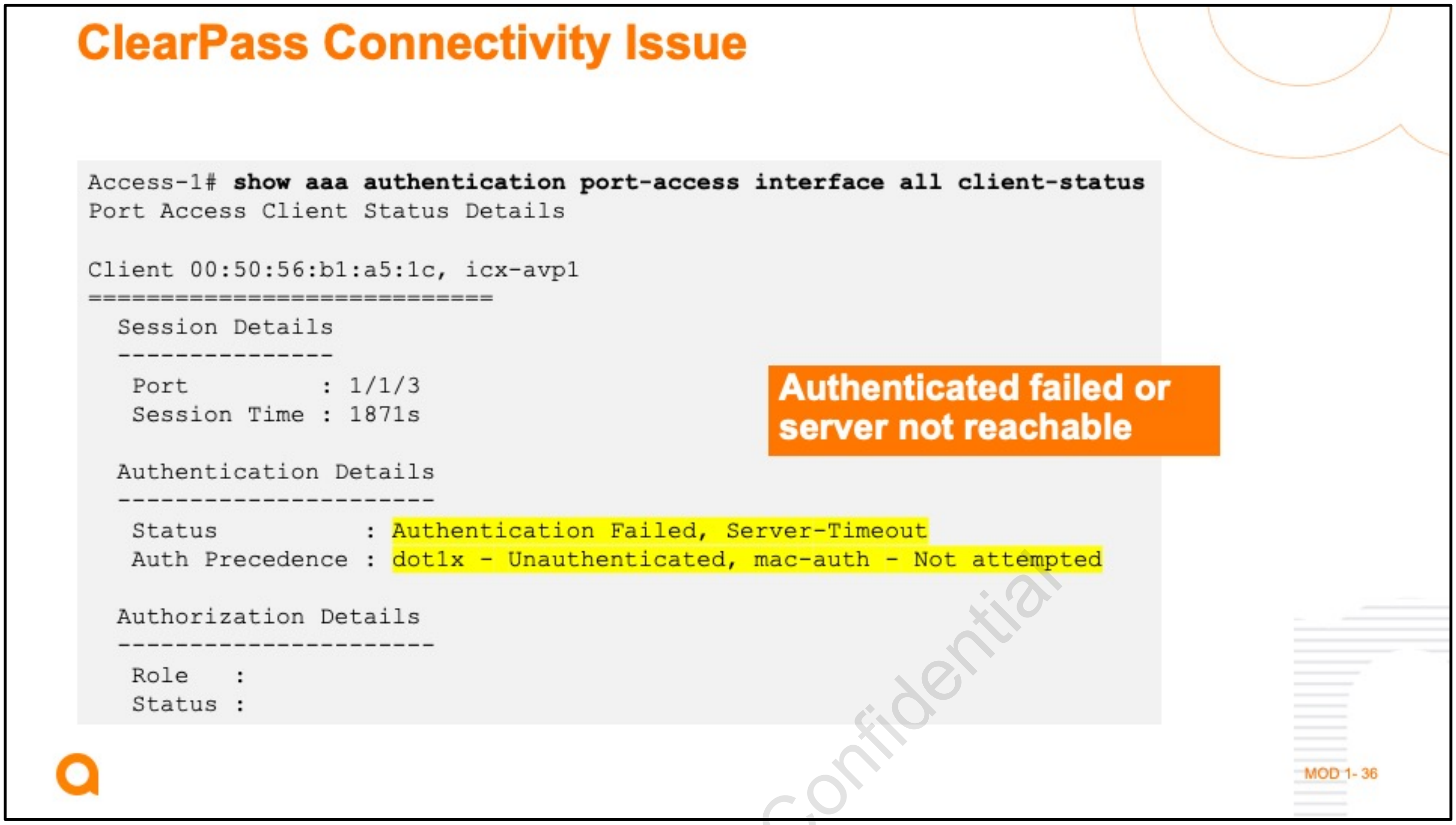
After switching over to the backup controller, the switch continues to check whether the primary controller is available every 30 seconds. If it re-establishes communication with the primary controller, it moves its tunnel back to that controller.

Again, MC administrators might choose an active-standby or active-active deployment. Communicate with them to determine whether all switches will use the same primary controller IP and backup controller IP (active-standby) or whether you should alternate these roles on different switches (active-active). Often, the MCs are deployed in different locations, and each switch uses the local MC as its primary MC and the remote MC as its backup.

Verification

MOD 1- 35

This last section examines some commands you can use to verify the configuration and operation of dynamic segmentation on your AOS-CX switch.

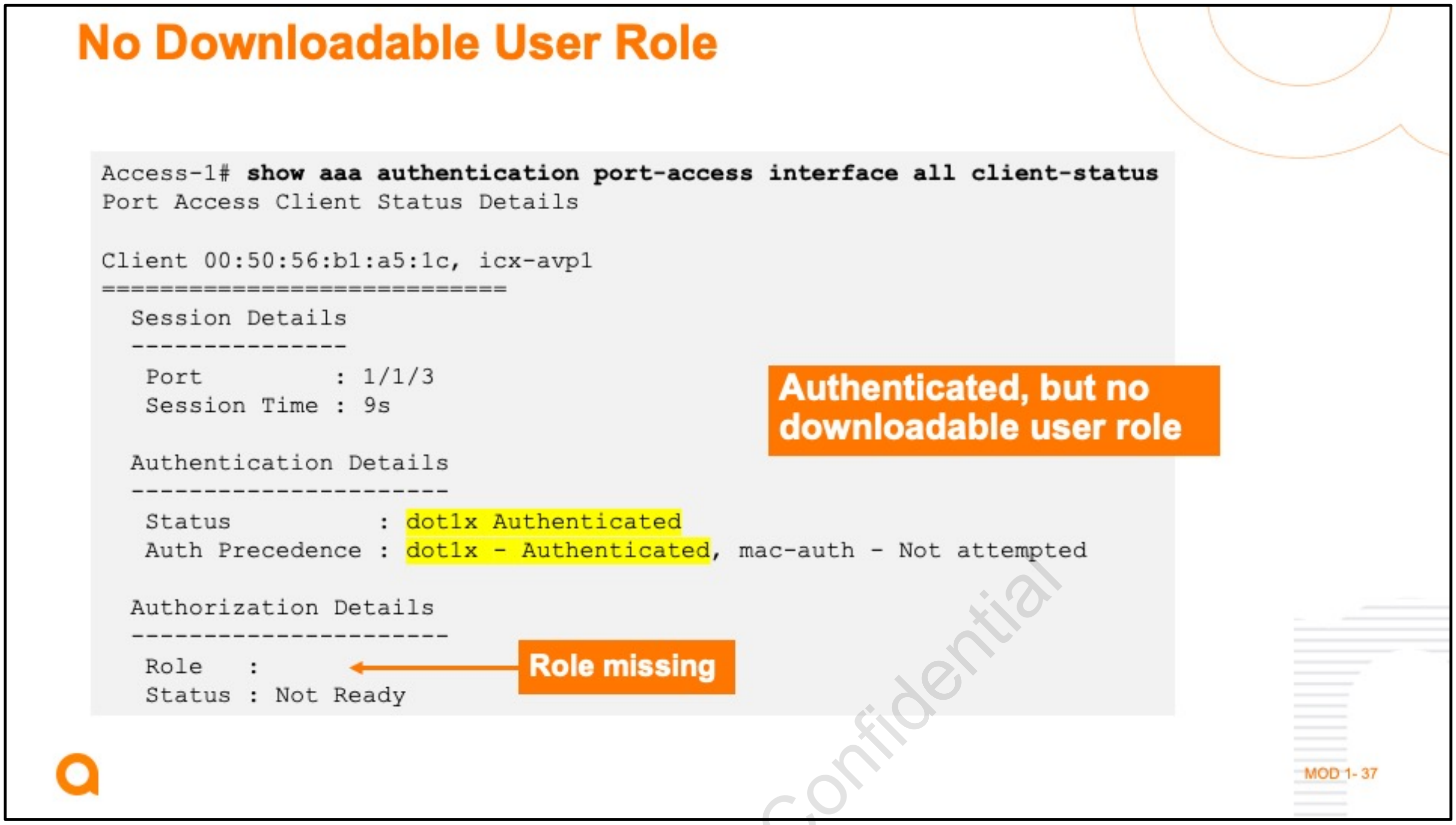


ClearPass Connectivity Issue

```
Access-1# show aaa authentication port-access interface all client-status
Port Access Client Status Details

Client 00:50:56:b1:a5:1c, icx-avp1
============================
  Session Details
  ---------------
   Port         : 1/1/3
   Session Time : 1871s

  Authentication Details
  ----------------------
   Status          : Authentication Failed, Server-Timeout
   Auth Precedence : dot1x - Unauthenticated, mac-auth - Not attempted

  Authorization Details
  ----------------------
   Role   :
   Status :
```

Authenticated failed or server not reachable

The show aaa authentication port-access interface all client-status command allows you to see the status of a client(s) on a port(s), including their authentication status and assigned user role in any. This example shows an issue contacting the ClearPass server. Mostly commonly, certificate errors and the ClearPass username and password being misconfigured cause this issue.



No Downloadable User Role

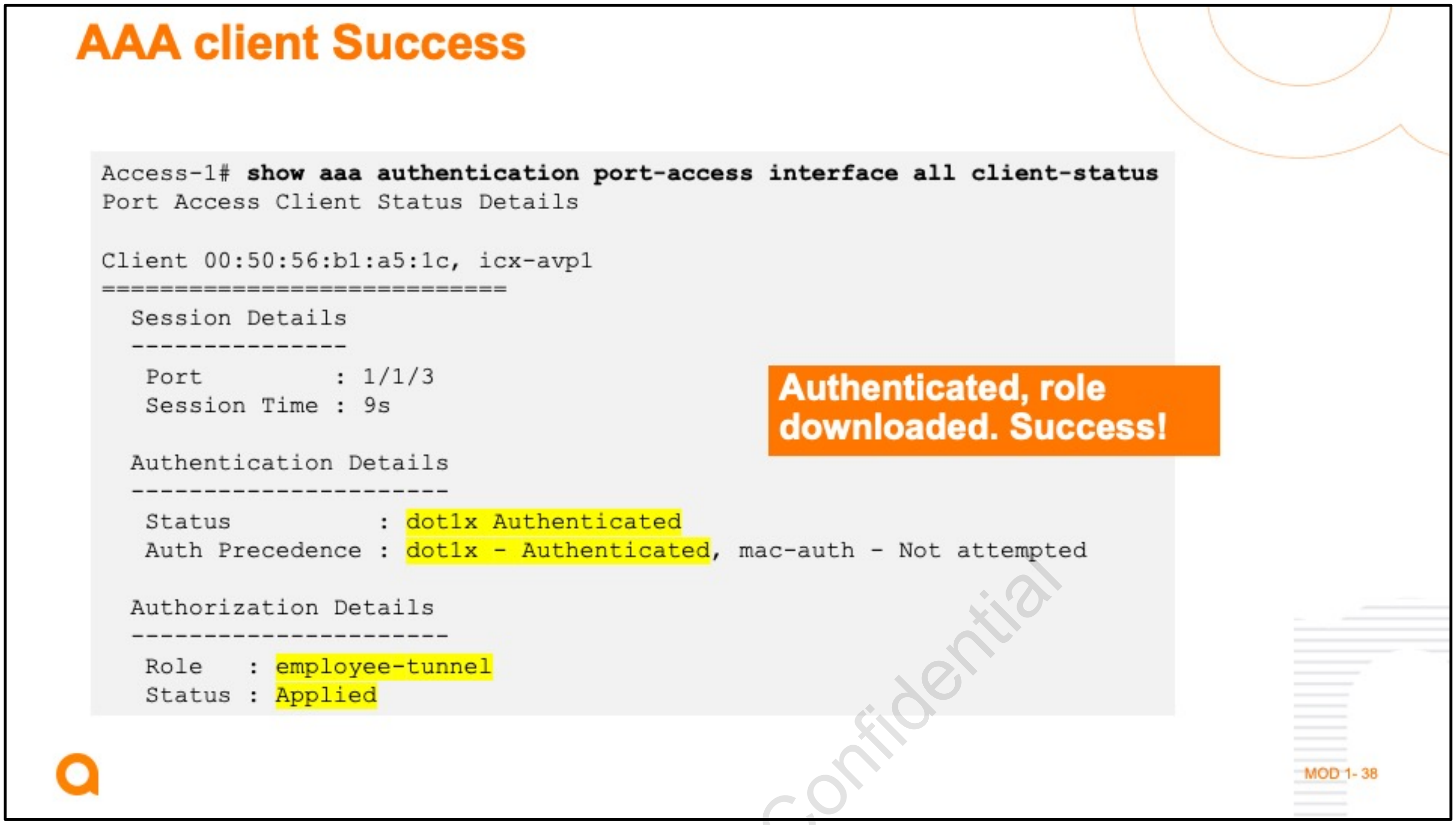
```
Access-1# show aaa authentication port-access interface all client-status
Port Access Client Status Details

Client 00:50:56:b1:a5:1c, icx-avp1
============================
  Session Details
  ---------------
   Port         : 1/1/3
   Session Time : 9s

  Authentication Details
  ----------------------
   Status          : dot1x Authenticated
   Auth Precedence : dot1x - Authenticated, mac-auth - Not attempted

  Authorization Details
  ----------------------
   Role   :
   Status : Not Ready
```

Authenticated, but no downloadable user role

Role missing

MOD 1- 37

This example shows a status of “Not Ready”. In this example, ClearPass passed down the user role called “Authenticated”, but this doesn’t exist in a port-access role command.



AAA client Success

```
Access-1# show aaa authentication port-access interface all client-status
Port Access Client Status Details

Client 00:50:56:b1:a5:1c, icx-avp1
============================
  Session Details
  ---------------
   Port         : 1/1/3
   Session Time : 9s

  Authentication Details
  ----------------------
   Status          : dot1x Authenticated
   Auth Precedence : dot1x - Authenticated, mac-auth - Not attempted

  Authorization Details
  ----------------------
   Role   : employee-tunnel
   Status : Applied
```

Authenticated, role downloaded. Success!

MOD 1- 38

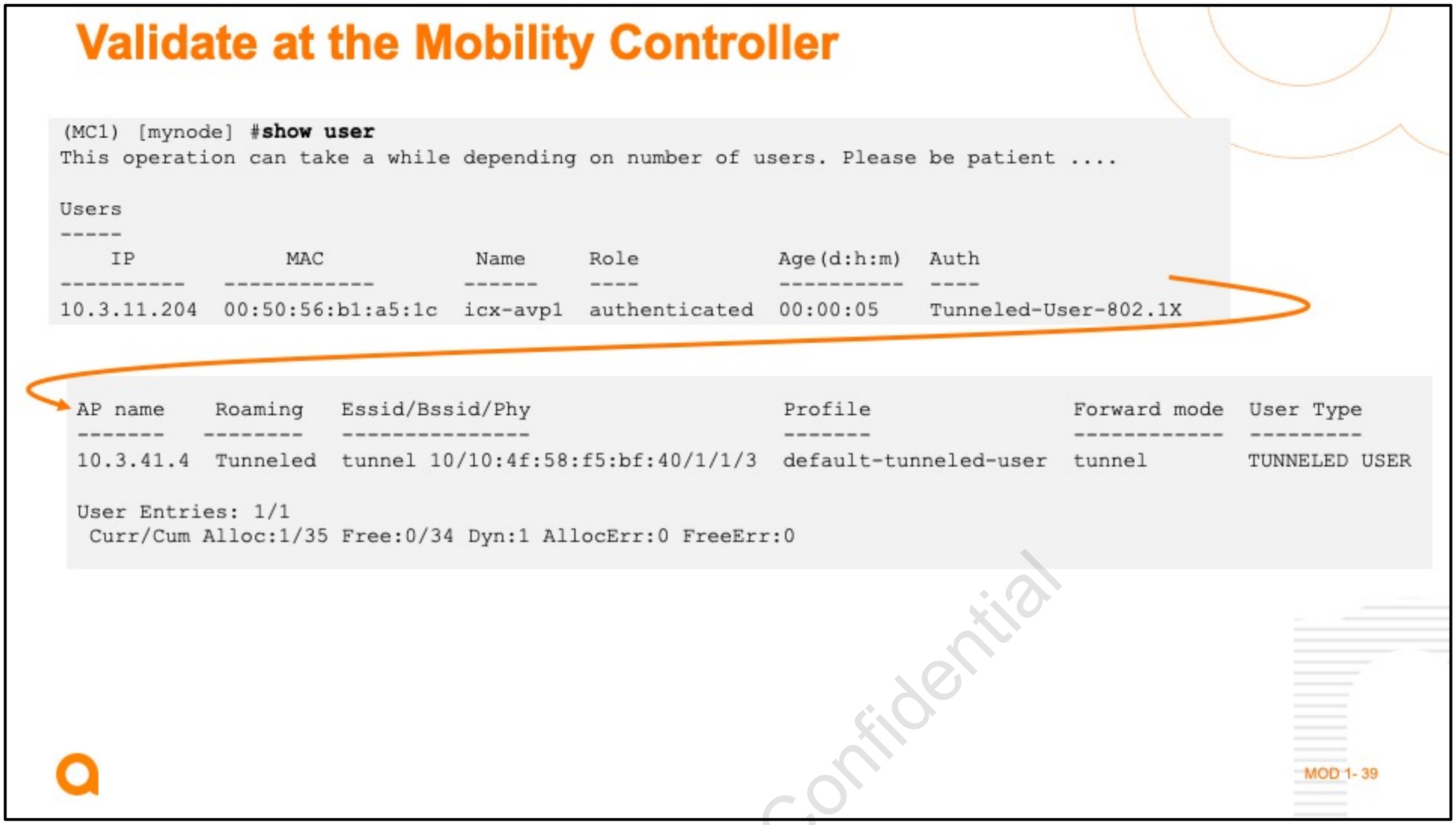
This example shows a Status of “Applied”, indicating that everything is working normal and that the Authenticated user role is applied to the client connected to port 1/1/3.

Applying the profile makes the port function like a controlled Aruba AP. The switch no longer forwards traffic that arrives on the ports locally as it normally would. Instead it uses the GRE tunnel to send all the traffic to the IP address that you configured in the tunnel server profile. The tunnel acts at Layer 2. It encapsulates not just the IP payload but also the MAC header, and it forwards all traffic including broadcasts and multicasts.

The main difference between the tunneled-node and AP-based tunnels is that tunneled-node encapsulates 802.3 instead of 802.11 frames. Note also that often wireless traffic is encrypted; however, wired traffic is not encrypted, and GRE does not provide encryption either. However, the wired network infrastructure is typically trusted, so encryption is not required.

Note that the switch continues to function as normal for all traffic that arrives on the non-tunneled-node ports.

Important: Tunneled-node ports are intended for wired endpoints. You should avoid plugging controlled Aruba APs into wired tunneled ports because these APs are already fulfilling similar functions on their own. You would be causing the network to encapsulate wireless clients' traffic twice. Lab tests show that the tunneled-node will still operate properly, but the double encapsulation can create performance and resource issues.



Validate at the Mobility Controller

```
(MC1) [mynode] #show user
This operation can take a while depending on number of users. Please be patient ....

Users
-----
    IP           MAC              Name    Role          Age(d:h:m)  Auth
----------  ------------      ------  ----          ----------  ----
10.3.11.204  00:50:56:b1:a5:1c  icx-avp1  authenticated  00:00:05    Tunneled-User-802.1X
```

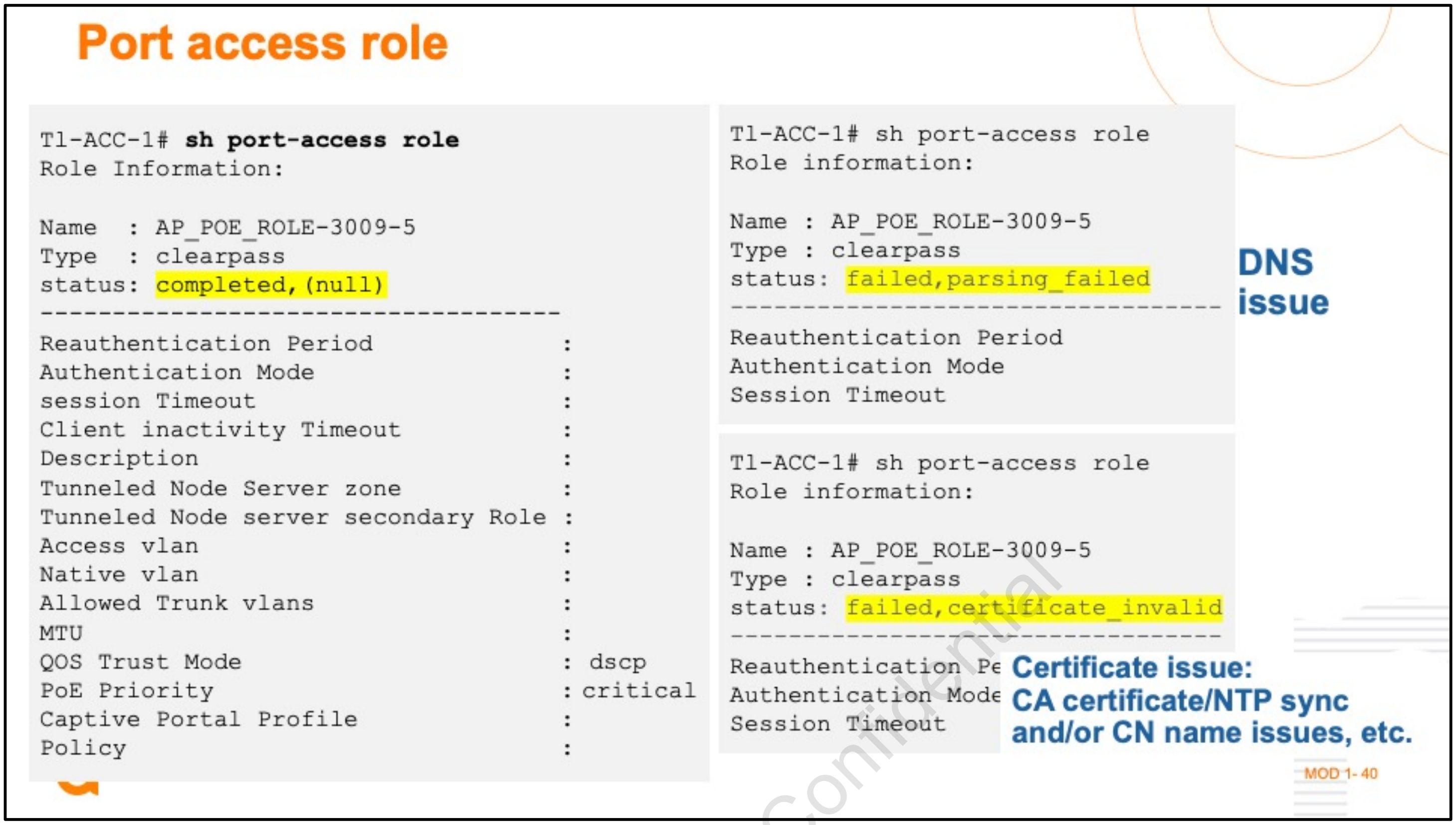
```
AP name    Roaming    Essid/Bssid/Phy                   Profile                  Forward mode  User Type
-------   --------    ---------------                   -------                  ------------  ---------
10.3.41.4  Tunneled   tunnel 10/10:4f:58:f5:bf:40/1/1/3  default-tunneled-user  tunnel        TUNNELED USER

User Entries: 1/1
 Curr/Cum Alloc:1/35 Free:0/34 Dyn:1 AllocErr:0 FreeErr:0
```

MOD 1- 39

This example shows how to validate connectivity from the Aruba Mobility Controller with the show user command. From left to right, you see the end user's IP and MAC address, name and role, time of connectivity, and authentication.

The next columns start with the source IP address of the Access switch, under the heading "AP name". You see that this is a tunneled connection, along with tunnel information – switch MAC address and port number. Then you see the assigned profile, forwarding mode, the user type – TUNNELED USER.

Note: The columns labeled VPN link, Type, and Host Name are not shown, to focus more relevant information.



Port access role

```
T1-ACC-1# sh port-access role
Role Information:

Name   : AP_POE_ROLE-3009-5
Type   : clearpass
status: completed,(null)
------------------------------------
Reauthentication Period                 :
Authentication Mode                     :
session Timeout                         :
Client inactivity Timeout               :
Description                             :
Tunneled Node Server zone               :
Tunneled Node server secondary Role :
Access vlan                             :
Native vlan                             :
Allowed Trunk vlans                     :
MTU                                     :
QOS Trust Mode                          : dscp
PoE Priority                            : critical
Captive Portal Profile                  :
Policy                                  :
```

```
T1-ACC-1# sh port-access role
Role information:

Name : AP_POE_ROLE-3009-5
Type : clearpass
status: failed,parsing_failed
------------------------------------
Reauthentication Period
Authentication Mode
Session Timeout
```

DNS issue

```
T1-ACC-1# sh port-access role
Role information:

Name : AP_POE_ROLE-3009-5
Type : clearpass
status: failed,certificate_invalid
------------------------------------
Reauthentication Pe
Authentication Mode
Session Timeout
```

Certificate issue: CA certificate/NTP sync and/or CN name issues, etc.

MOD 1- 40

The show port-access role command can be used to verify the connectivity to the ClearPass server when using DUR. The left-hand example shows a successful connection and inter-operation with the ClearPass server: the Status is completed. The top-right example shows a parsing_failed status, typically indicative of either a DNS or network connectivity issue. The bottom-right example shows an certificate error, which could be caused by a number of things:

Time issue (use NTP to ensure the correct time)

The CN name of the ClearPass certificate is not defined on the switch correctly (or at all)

The certificate has expired

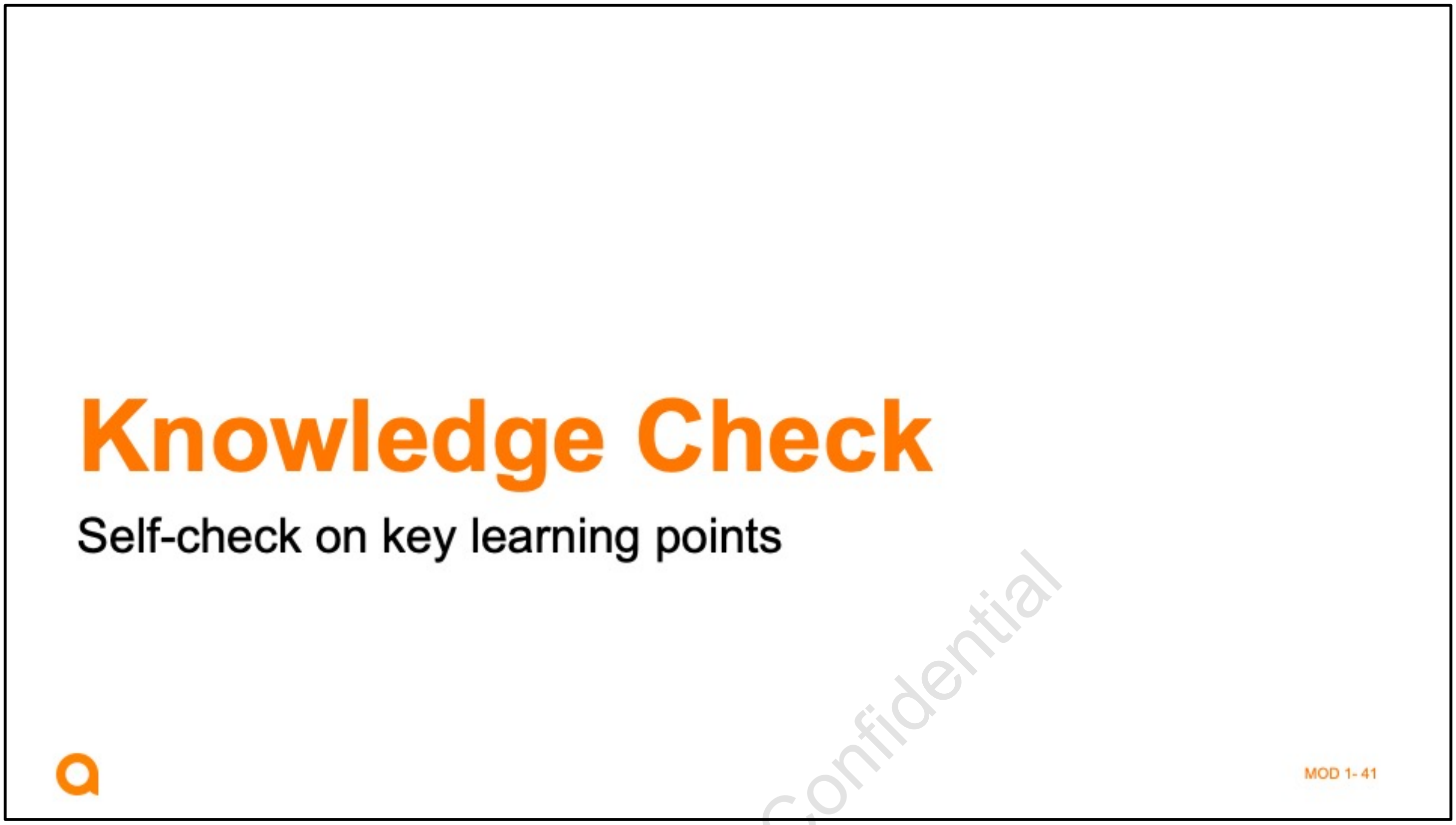
Controller commands

To verify the connectivity from the controller side, use the show tunneled-node-mgr tunneled-nodes command.

To see the users that are being tunneled from the switch, on the MC use the show tunneled-node-mgr-tunneled-users command.

To see the log messages related to tunneled node on the MCs, use the show tunneled-node-mgr trace-buf command.

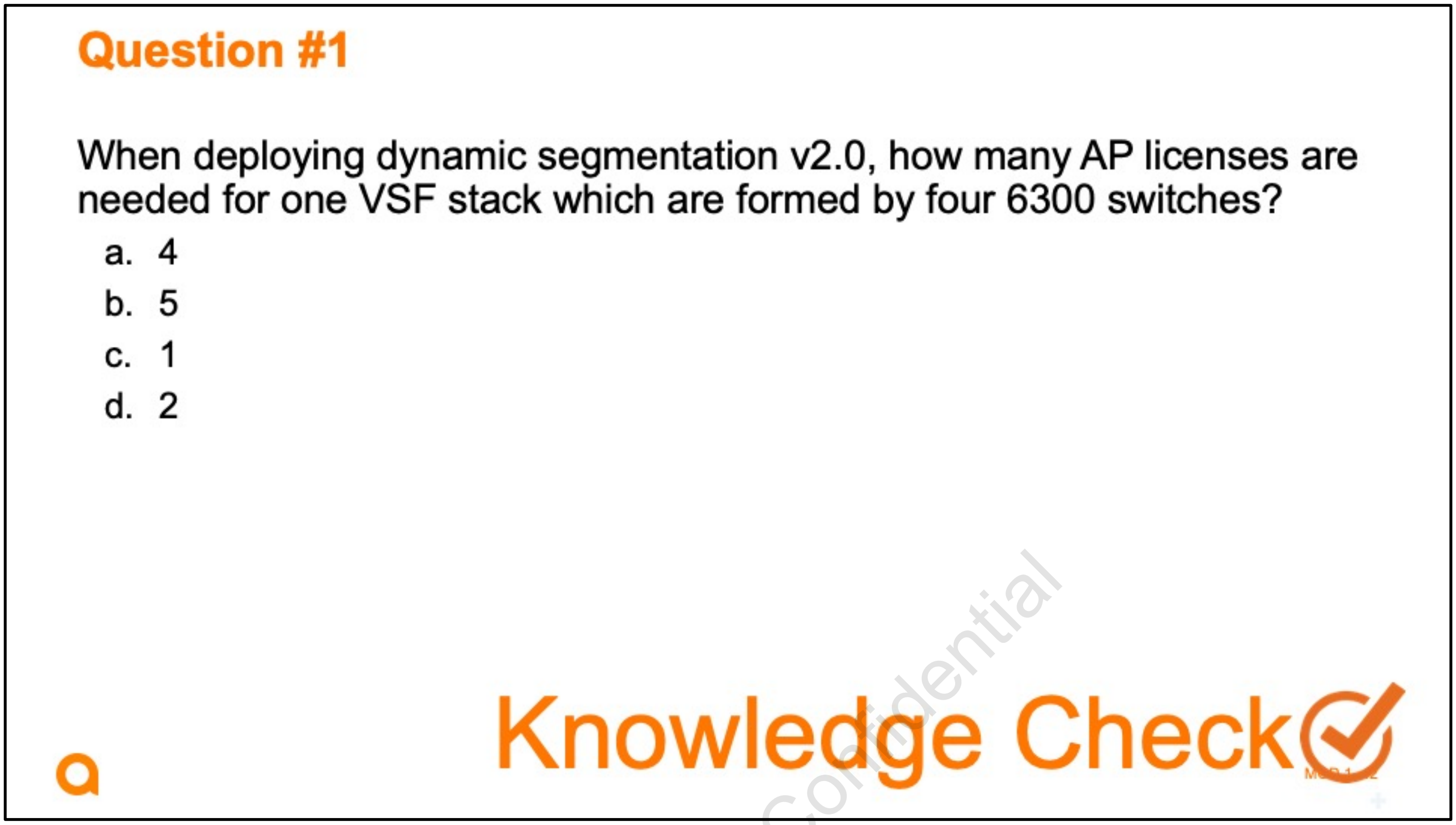
To see the tunneled connections in the MC’s state table, use the show datapath command.



Knowledge Check

Self-check on key learning points

MOD 1- 41



Question #1

When deploying dynamic segmentation v2.0, how many AP licenses are needed for one VSF stack which are formed by four 6300 switches?

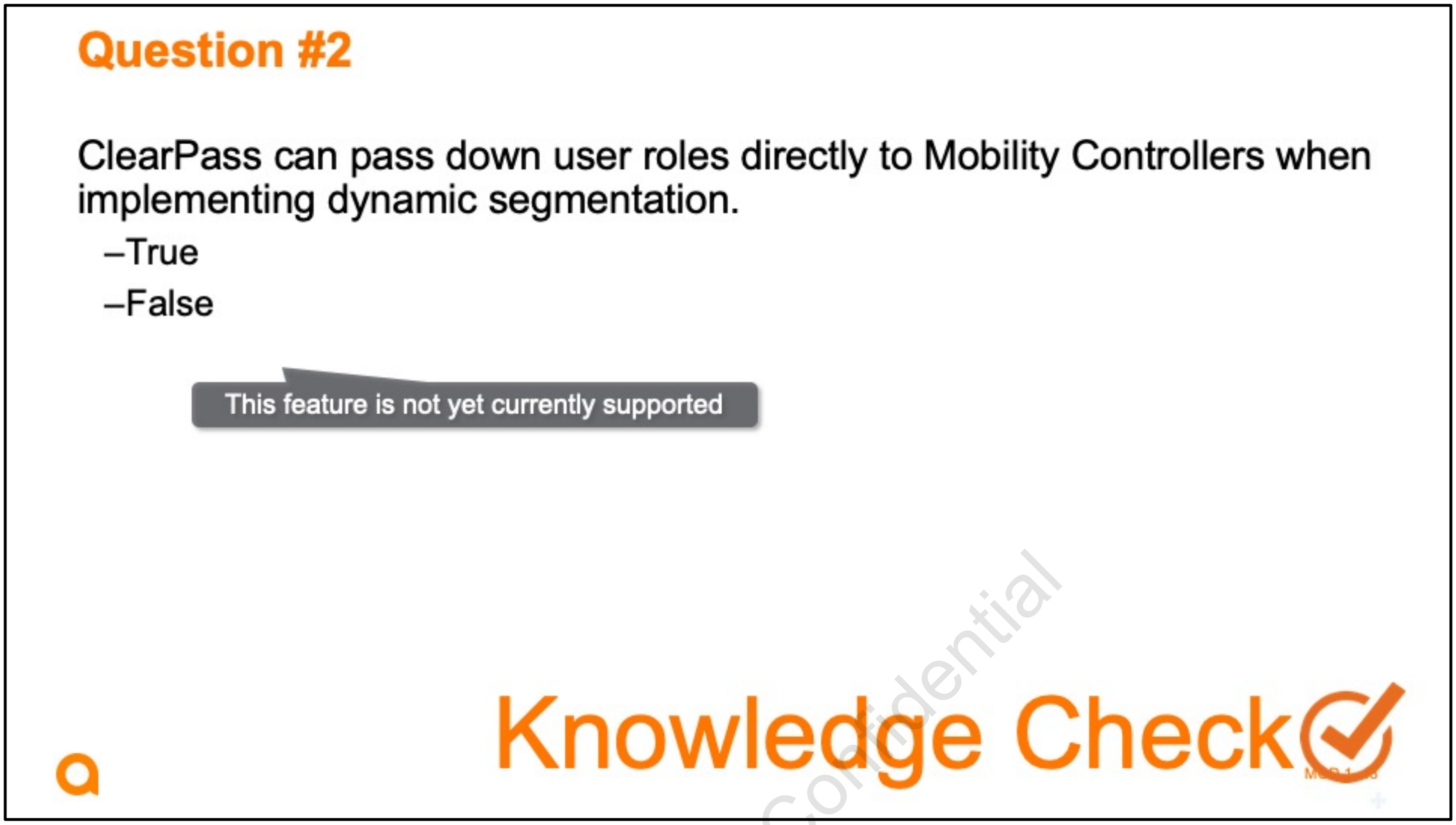
a. 4

b. 5

c. 1

d. 2

Knowledge Check



Question #2

ClearPass can pass down user roles directly to Mobility Controllers when implementing dynamic segmentation.

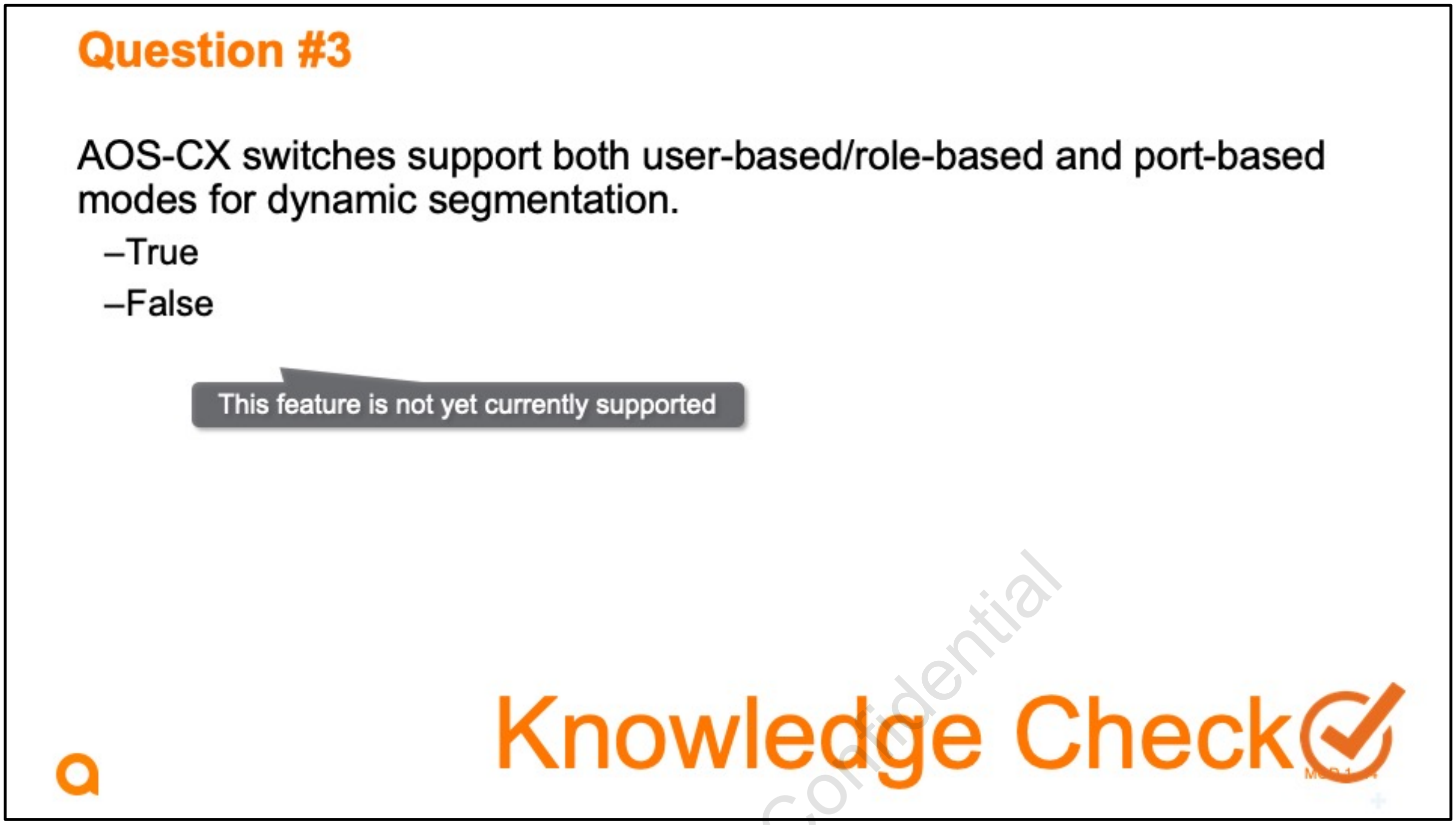
–True

–False

This feature is not yet currently supported

Knowledge Check

that text bubble in the figure “This feature is not yet currently supported” should appear when the correct answer appears.



Question #3

AOS-CX switches support both user-based/role-based and port-based modes for dynamic segmentation.

–True

–False

This feature is not yet currently supported

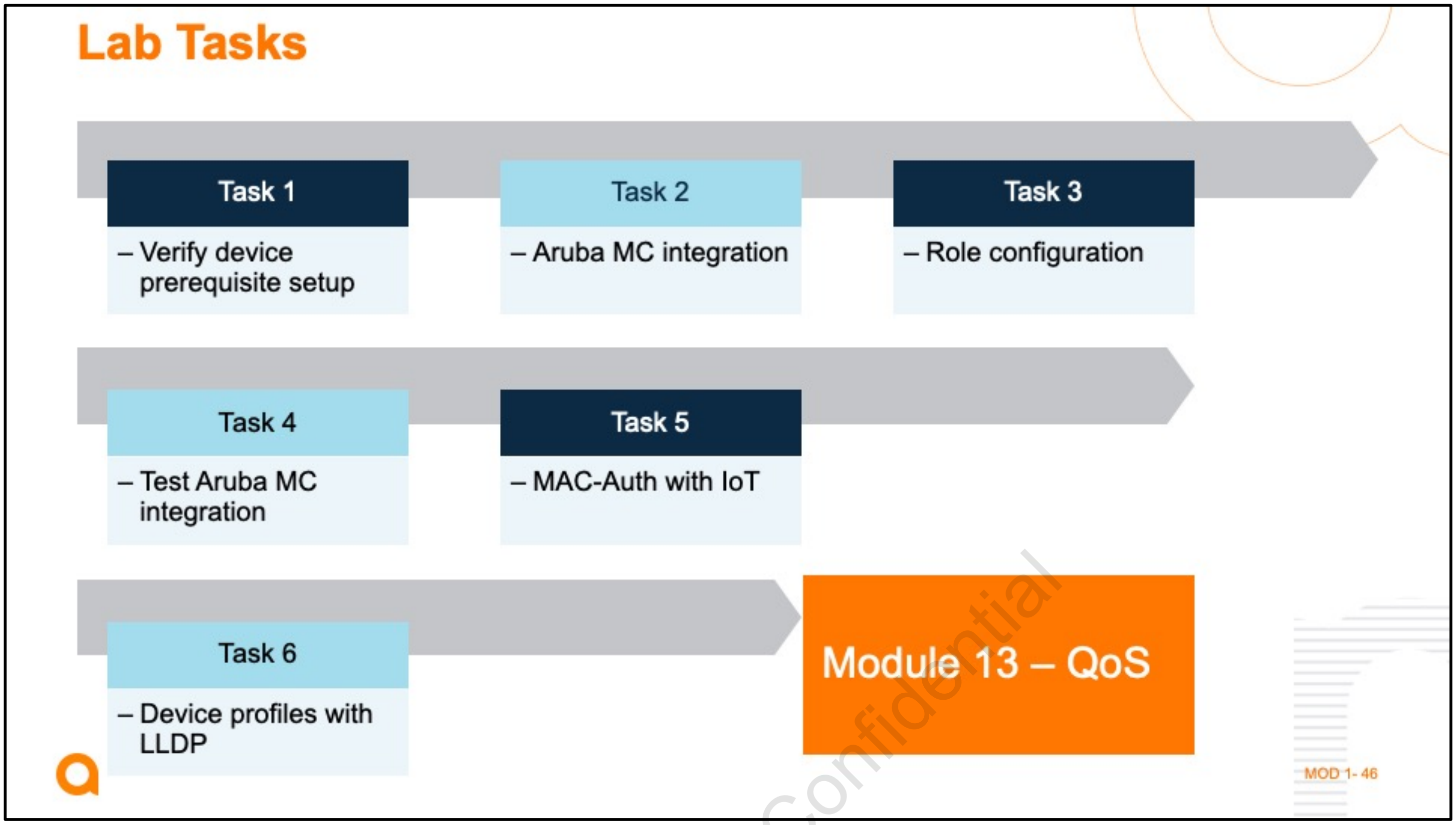
Knowledge Check

As always, any text bubble in the figure should appear when the correct answer appears, by adding it to the “feedback” for the question

Lab Activity

Lab 6.3

It time for a lab activity.



Lab Tasks

Task 1

– Verify device prerequisite setup

Task 2

– Aruba MC integration

Task 3

– Role configuration

Task 4

– Test Aruba MC integration

Task 5

– MAC-Auth with IoT

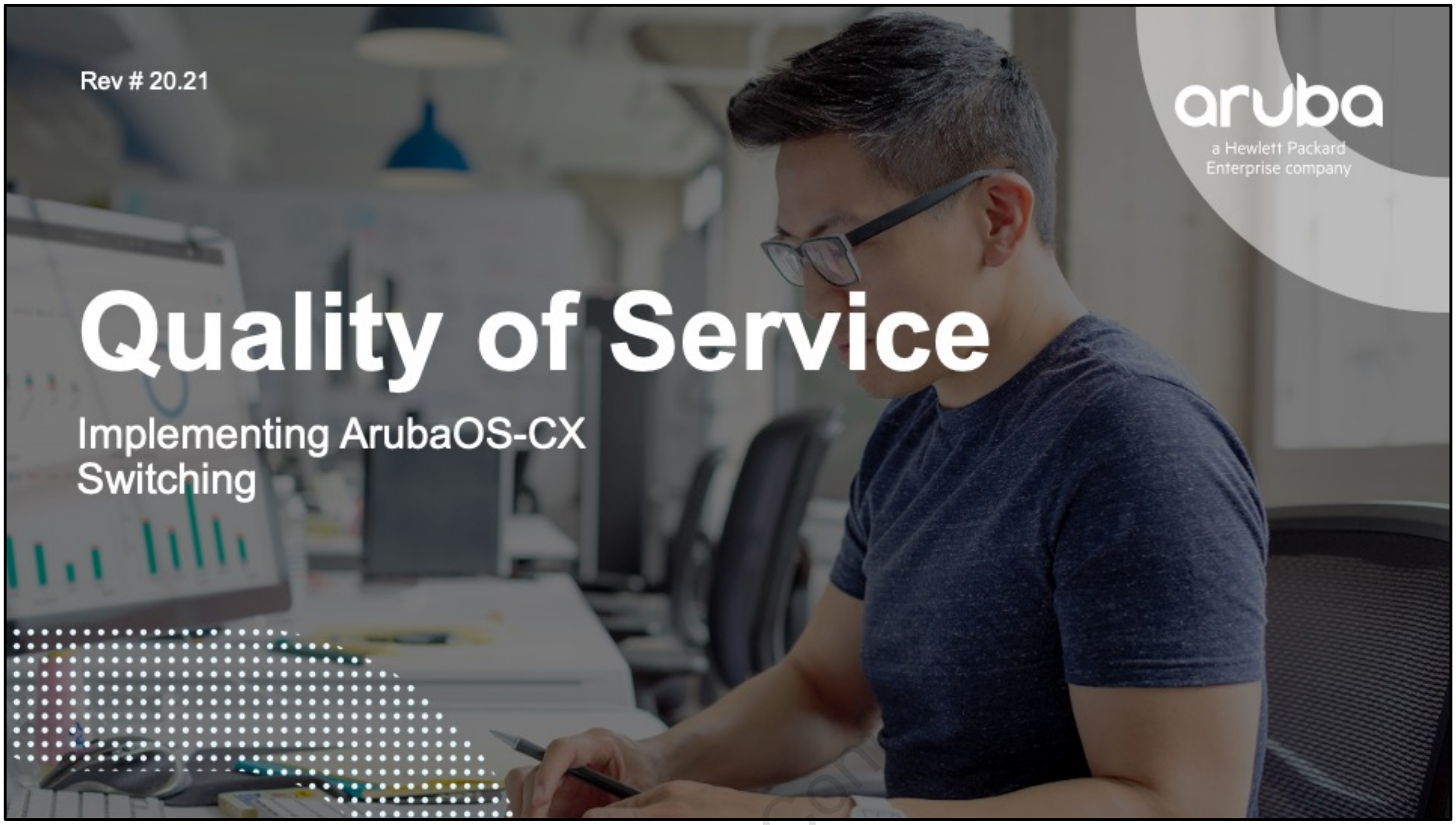
Task 6

– Device profiles with LLDP

Module 13 – QoS

MOD 1- 46

The figure provides a brief review of lab tasks. Please see your lab guide for details. When you are ready, please continue with Module 13 – QoS.

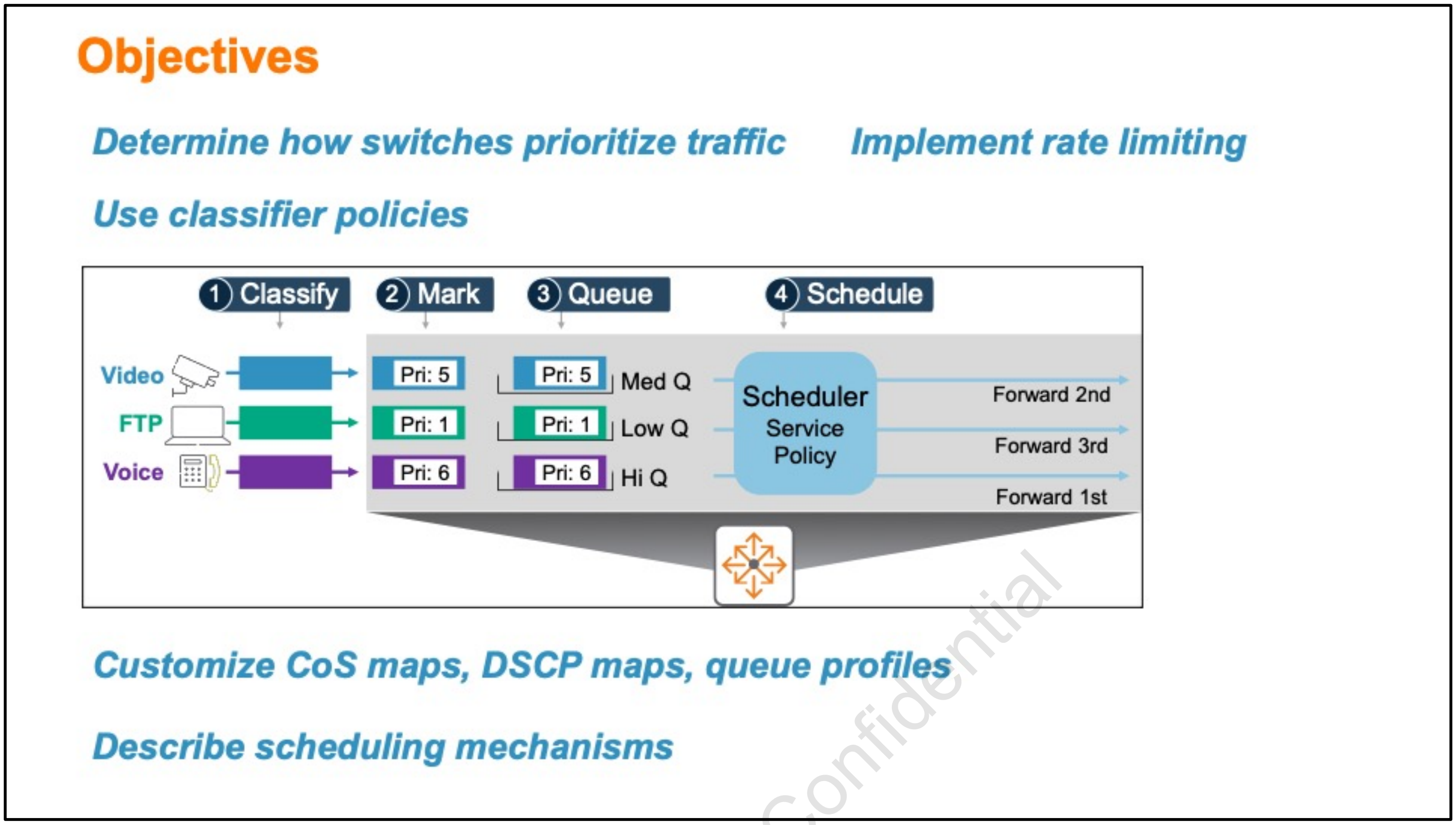


Rev # 20.21

Quality of Service

Implementing ArubaOS-CX Switching

Hello everyone! This is QoS.



Objectives

Determine how switches prioritize traffic

Implement rate limiting

Use classifier policies

Customize CoS maps, DSCP maps, queue profiles

Describe scheduling mechanisms

After completing this module, you should be able to:

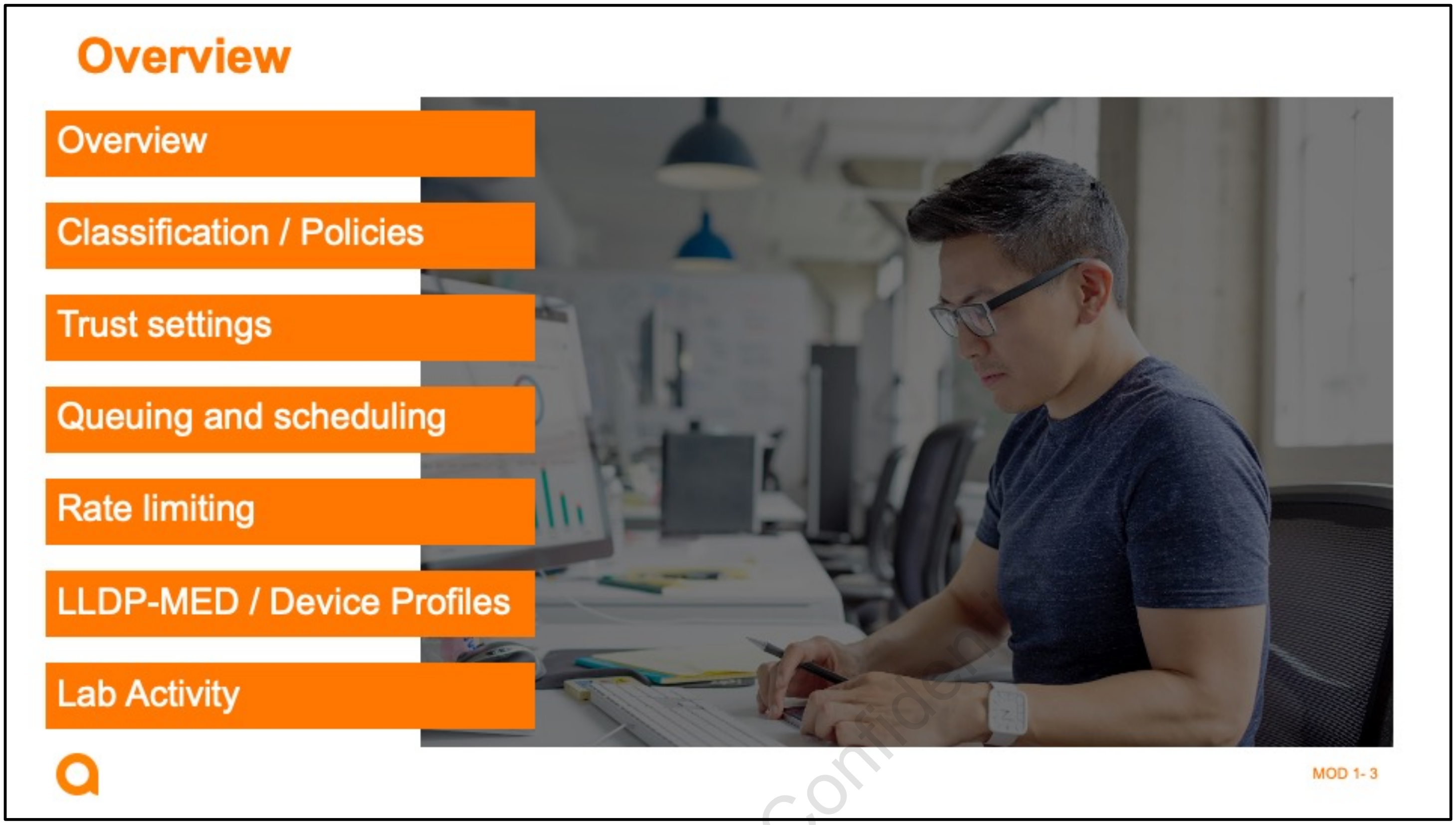
Determine how an AOS-CX switch will prioritize traffic based on incoming priority value and QoS trust settings

Implement rate limiting

Use classifier-based policies to assign traffic a new priority based on customer requirements

Customize CoS maps, DSCP maps, and queue profiles

Understand how scheduling mechanisms differ and select an appropriate scheduling profile



Overview

- Overview
- Classification / Policies
- Trust settings
- Queuing and scheduling
- Rate limiting
- LLDP-MED / Device Profiles
- Lab Activity

MOD 1- 3

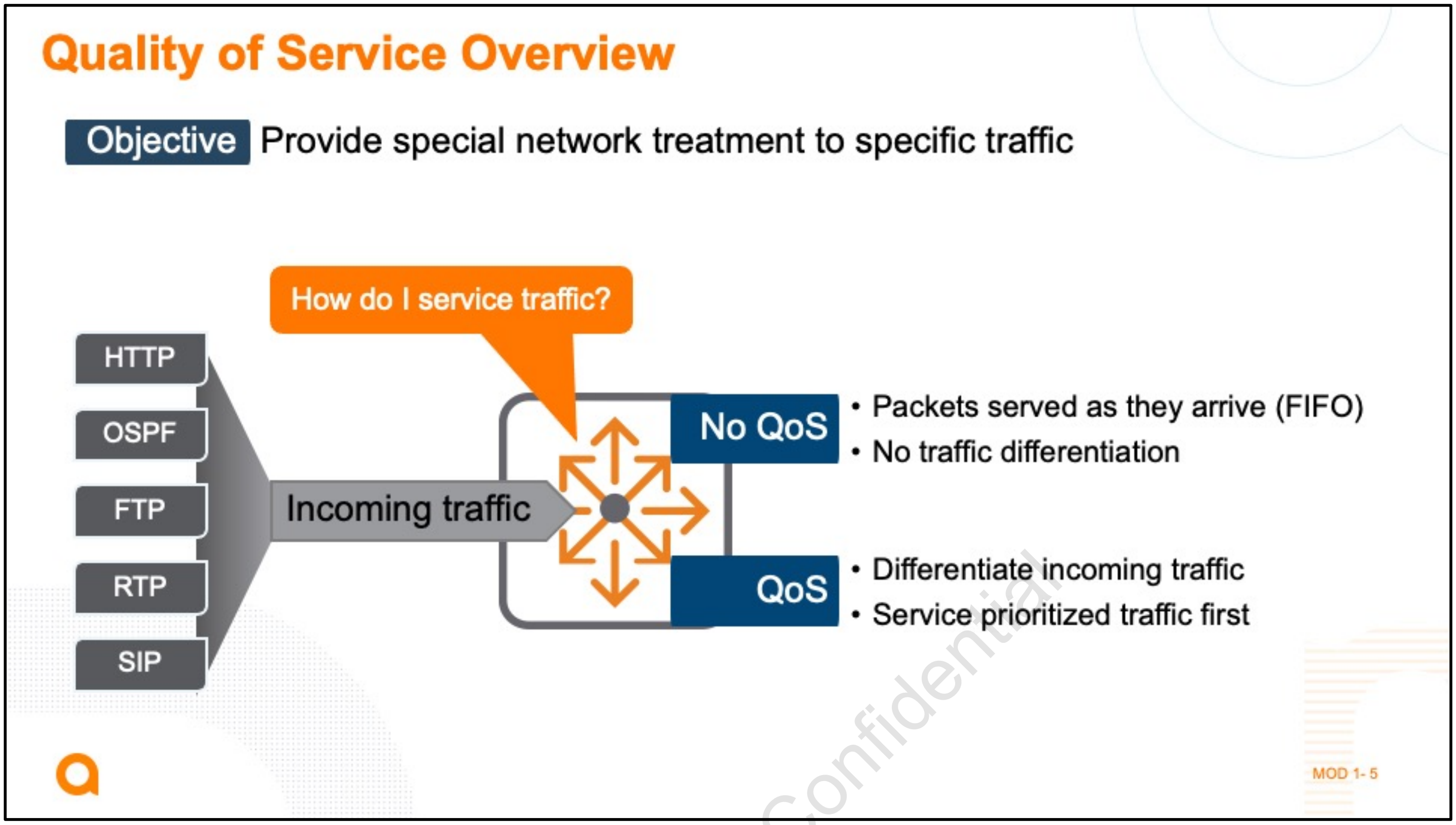
In this module you will learn how to implement Quality of Service (QoS) mechanisms that improve the experience of users running a variety of network applications from traditional ones to voice and video applications. The module teaches you how to implement the many sophisticated QoS features on AOS-CX switches.

You begin by learning about QoS objectives and how AOS-CX switches use a variety of mechanisms to classify and service traffic.

Then you dive into each of those mechanisms in more detail - trust settings, queuing, scheduling, rate limiting, and device profiles…all wrapped up with a lab activity.

Overview

We begin with an overview.



Quality of Service Overview

Objective Provide special network treatment to specific traffic

No QoS
- Packets served as they arrive (FIFO)
- No traffic differentiation

QoS
- Differentiate incoming traffic
- Service prioritized traffic first

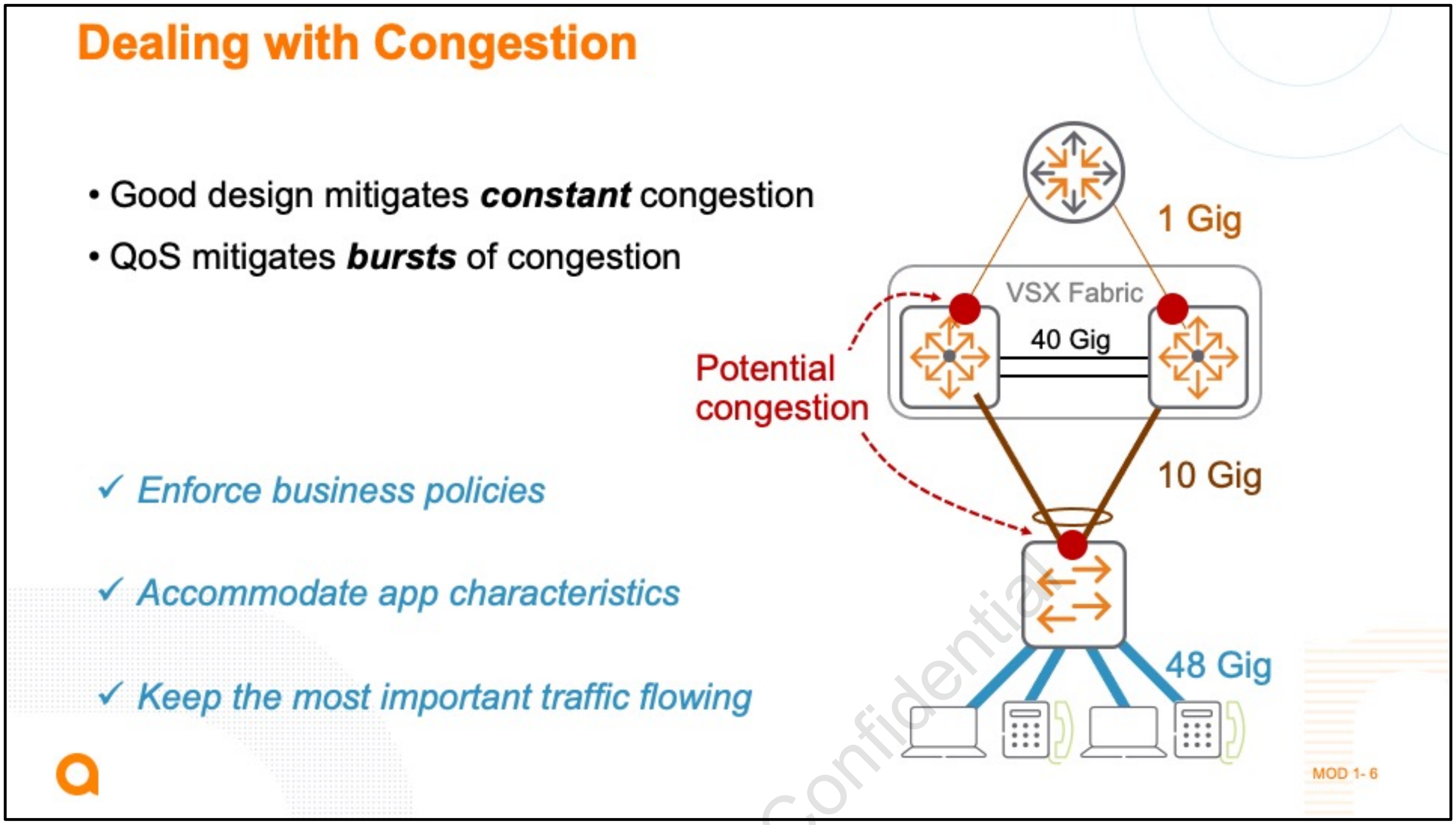
MOD 1- 5

Quality of Service (QoS) enables networks to provide better or “special” service to a set of users/applications, perhaps to the detriment of other users/applications. You make the best possible use of available bandwidth.

Different traffic types compete for the same network resources such as bandwidth and fast processing. Without QoS, packets are served as they arrive, typically First-In, First Out (FIFO). Your network management traffic may be more important to you than some users watching videos during their lunch break, but there is not differentiation between this traffic.

On the highway, an ambulance gets higher priority than other vehicles. Likewise, QoS is not being fair when dividing the network resources, but rather selecting certain traffic for improved service, perhaps at the expense of other traffic. QoS does not make the road wider, it just decides who goes first and consequently who must wait.

Understand that QoS is not a standalone service or product, but rather a concept that supports a framework of protocols, techniques, and network attributes that span across the infrastructure.



Dealing with Congestion

- Good design mitigates ***constant*** congestion
- QoS mitigates ***bursts*** of congestion

✓ *Enforce business policies*

✓ *Accommodate app characteristics*

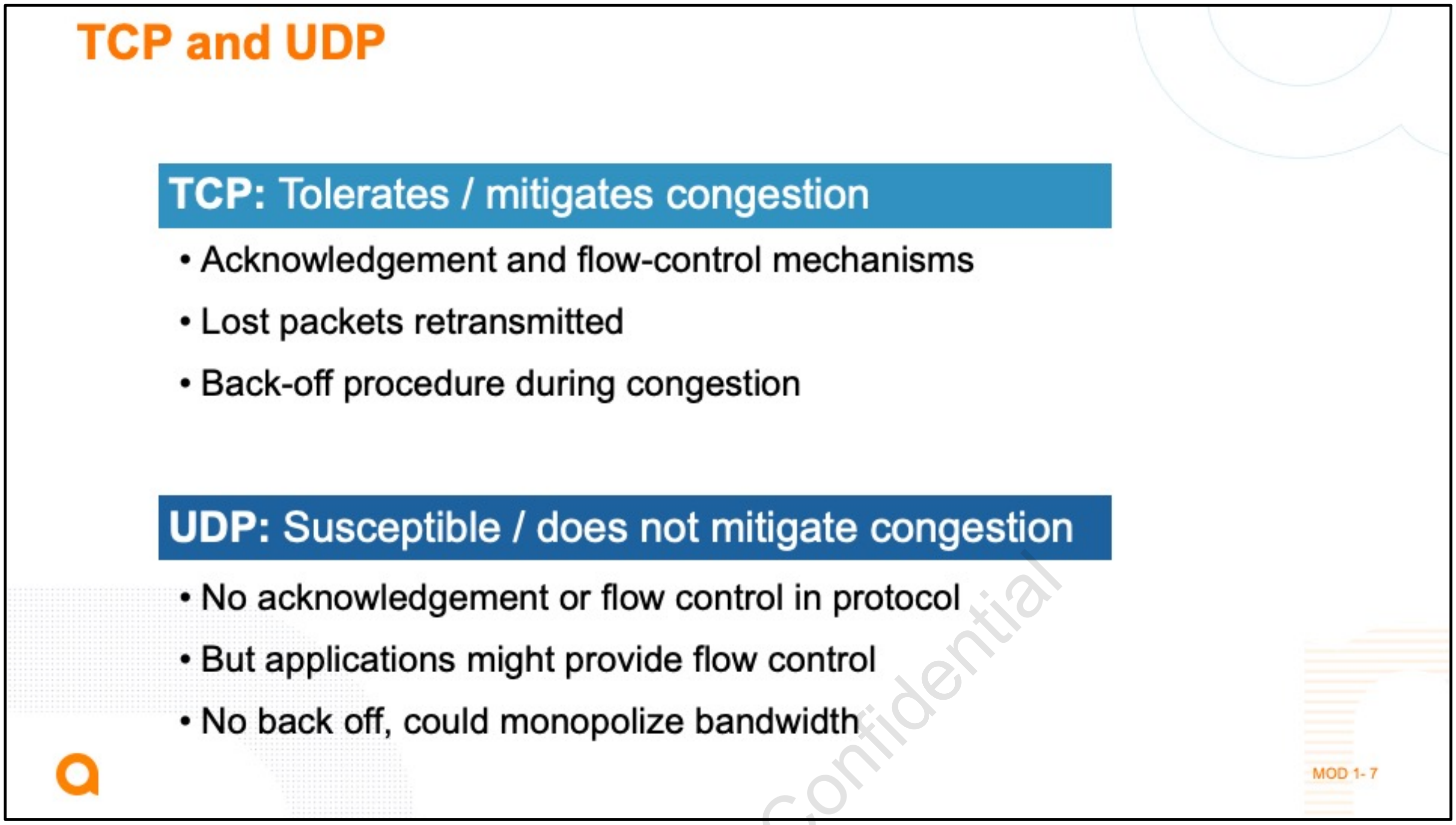
✓ *Keep the most important traffic flowing*

Even in a well-designed network, certain links may become oversubscribed – traffic arrives faster than it can be serviced. On a moment-to-moment basis, fluctuating traffic patterns and bursts of traffic can cause congestion. Lower-bandwidth links act as a bottleneck, and if traffic continues to arrive faster than the port can forward it, the port's buffers fill, and it drops traffic.

A properly designed network architecture can help to eliminate common sources of congestion. For example, a three-tier Access-Distribution-Core design can sometimes be flattened into a two-tier Access-Core design. This can reduce latency by eliminating the distribution layer. You should properly provision uplinks based on user and device needs. You should not design a network where several 48 Gigabit access links must all be serviced by two 10 Gigabit Access-to-Core links, which must then be serviced by two 1 Gigabit Core-to-WAN Edge links. As shown in the figure, this likely creates near constant congestion.

Still, even a well-designed network will experience occasional oversubscription and momentary bursts of congestion. Business-critical traffic might be dropped, and delay-sensitive applications might lose functionality. You should implement QoS mechanisms to protect sensitive traffic, ensure a good user experience.

QoS mechanisms help to enforce business policies that prioritize certain applications or network segments. You can customize how different traffic types are serviced. You accommodate unique application characteristics and their importance within the organization. QoS ensures uniform and efficient traffic handling, keeping the most important traffic moving at an acceptable speed, regardless of current bandwidth usage. It also enables you to control the priority settings of ingress traffic for each network device.



TCP and UDP

TCP: Tolerates / mitigates congestion

- Acknowledgement and flow-control mechanisms
- Lost packets retransmitted
- Back-off procedure during congestion

UDP: Susceptible / does not mitigate congestion

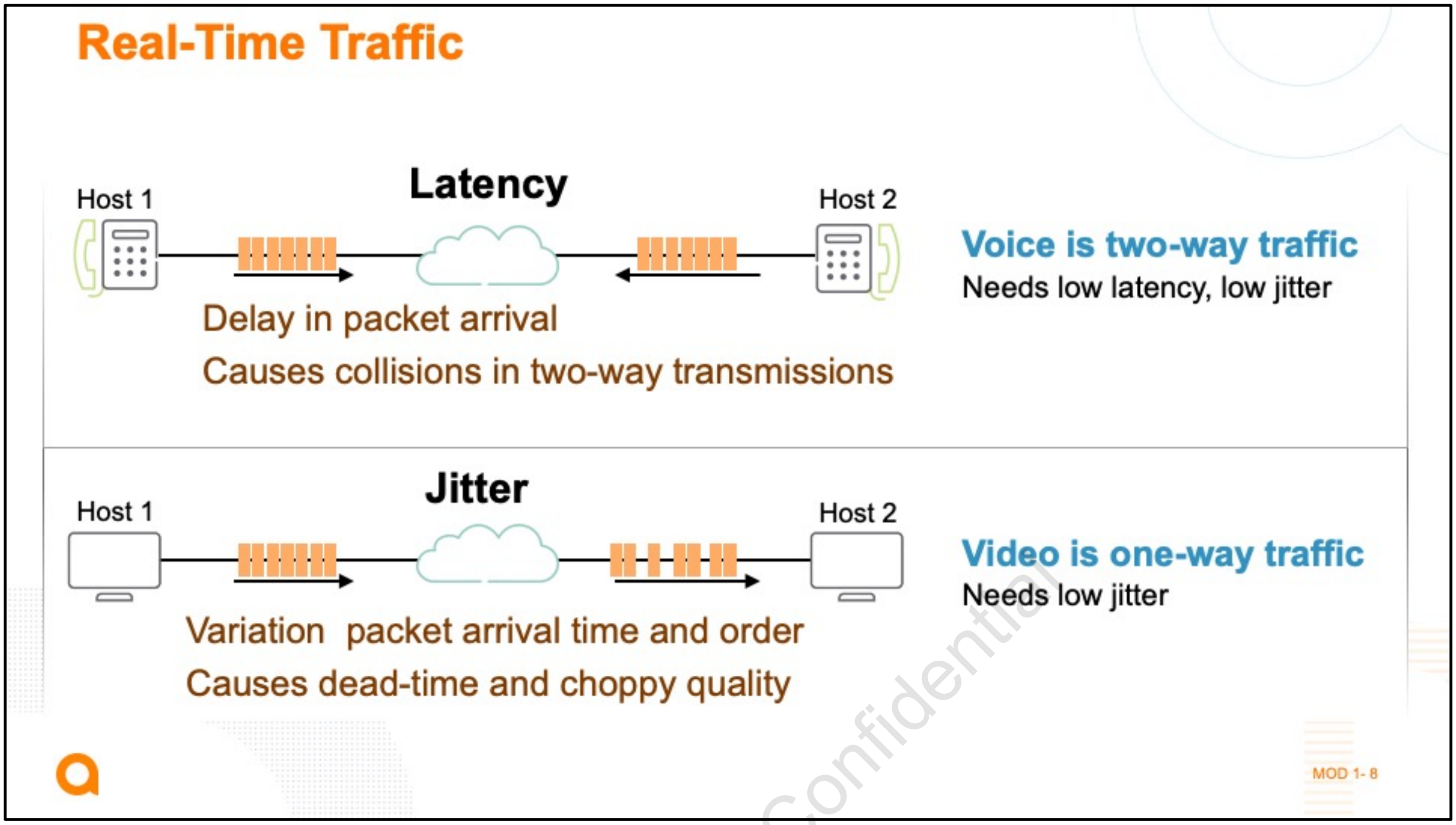
- No acknowledgement or flow control in protocol
- But applications might provide flow control
- No back off, could monopolize bandwidth

MOD 1- 7

In general, “best-effort” delivery is satisfactory for typical data traffic. Applications that use TCP for transport are less affected by packet loss because TCP provides acknowledgement and flow-control mechanisms. TCP detects and retransmits dropped packets. Each segment of a TCP/IP message includes sequence numbers that inform receivers of the proper order and number of packets in the transmission. The receiver in a TCP transmission acknowledges packets by returning messages that include the highest sequence number it has received. TCP applications include a “window” parameter that determines how many bytes can be outstanding and unacknowledged by the receiver at any given moment. This parameter limits the number of packets that can be acknowledged by a single control packet. Delay in delivery of packets or excessive packet loss prompts the sender and receiver to establish a smaller window size. Furthermore, TCP applications back off when network congestion is detected, enabling applications to share bandwidth.

UDP, on the other hand, does not provide acknowledgements or flow control. This reduced overhead can benefit delay-sensitive applications. Also, some applications include their own error-checking procedures, and do not need TCP to do it for them. However, in most cases, UDP applications are designed to minimize the need for retransmission and use UDP as a transport mechanism precisely because of its lower overhead and network efficiency. UDP also does not provide any back-off procedures.

Because UDP applications do not respond to congestion in any way, they can continue to contribute to congestion, which can also have drastic effects on time-sensitive real-time traffic such as voice and video.



Real-Time Traffic

Latency

Delay in packet arrival

Causes collisions in two-way transmissions

Voice is two-way traffic

Needs low latency, low jitter

Jitter

Variation packet arrival time and order

Causes dead-time and choppy quality

Video is one-way traffic

Needs low jitter

MOD 1- 8

Voice traffic is sensitive to delay, sometimes called latency, which is the amount of time that passes between the sending of a transmission and its arrival at the receiving station.

The figure shows small packets traveling from sender to receiver. In both examples, the packets are sent at a consistent rate.

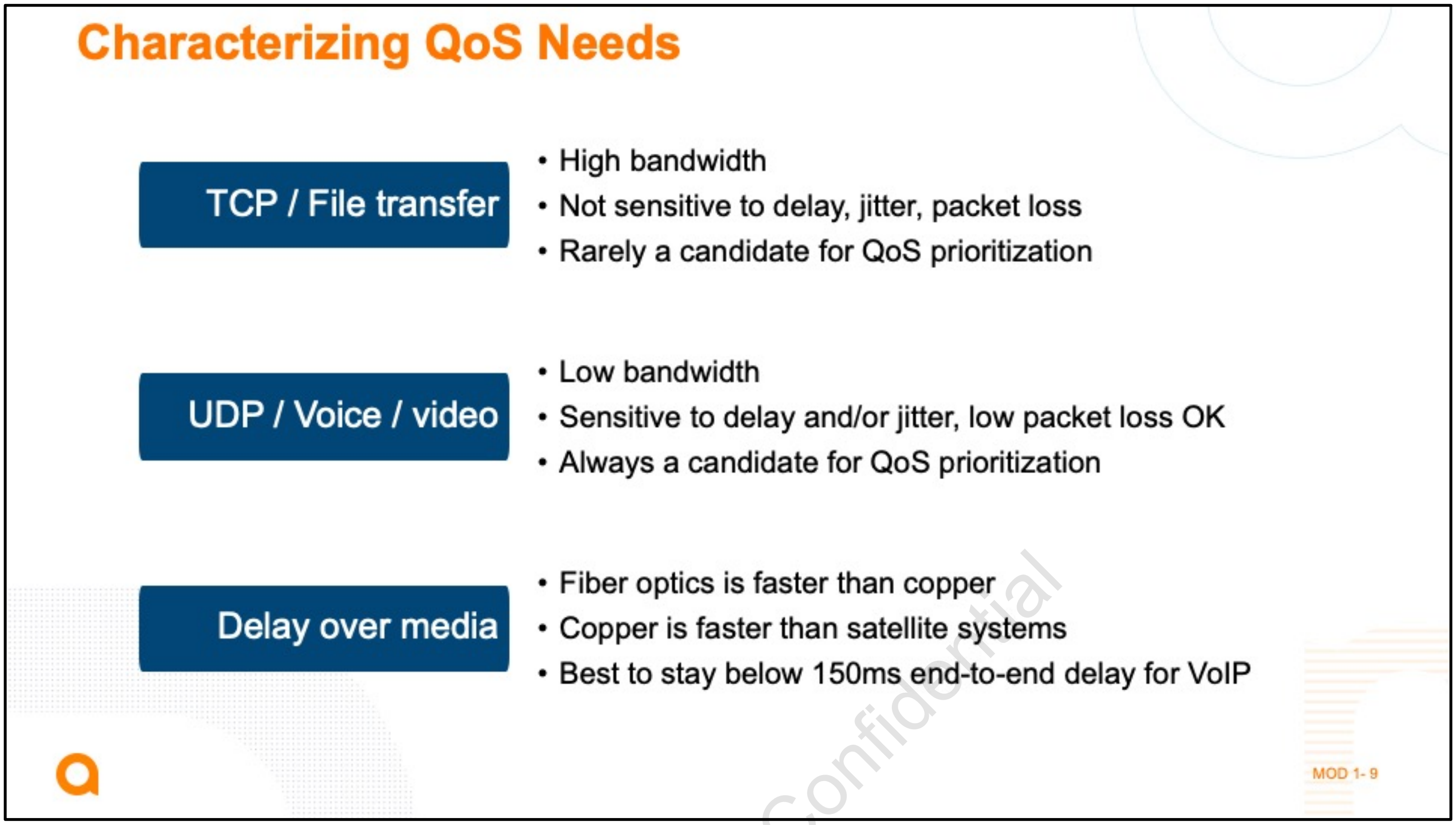
In the latency example, each device in the network cloud requires time to process each packet. Add these times to that required for packets to traverse each link in the cloud, and you get total latency. The time between packet sent from Host 1 to arrival at Host 2.

In the case of unidirectional traffic (as video traffic often is), this predictable, measurable time period is usually tolerable, regardless of its length. However, because multidirectional voice traffic is nearly always interactive, excessive delay can result in conversation collisions. Suppose that Host 1 begins transmitting in the belief that Host 2 is not transmitting. However, because Host 2 is still transmitting, bits from Host 2 begin arriving before Host 1 concludes its transmission. The conversations overlap.

Voice traffic is also sensitive to jitter, which is the variation in intervals between the arrivals of packets or, in other words, the difference between when a packet is expected and when it actually arrives. Video is also sensitive to jitter.

In the Jitter example, Host 1 sends data at a constant rate to Host 2. However, devices within the cloud that represents the network experience varying levels of congestion that result in some packets being delayed more than others. The outcome is a variation in the interval between packet arrival, which results in a choppy voice or video stream.

Jitter can have significant effect on real-time applications even if they require very small data streams. For instance, an IP telephone sends very small packets every 20-30 milliseconds (ms). For the receiver to properly interpret the stream, the packets must arrive at the same rate. If the interval between packets grows to 50ms, for instance, the transmission will be unsatisfactory for many users.



Characterizing QoS Needs

TCP / File transfer

- High bandwidth
- Not sensitive to delay, jitter, packet loss
- Rarely a candidate for QoS prioritization

UDP / Voice / video

- Low bandwidth
- Sensitive to delay and/or jitter, low packet loss OK
- Always a candidate for QoS prioritization

Delay over media

- Fiber optics is faster than copper
- Copper is faster than satellite systems
- Best to stay below 150ms end-to-end delay for VoIP

MOD 1- 9

All traffic does not require the same handling to receive an appropriate level of service. You should understand common applications and their characteristics, as relates to QoS. This is summarized in the figure - click on each row to explore.

TCP/File Transfer

Many TCP applications can tolerate some service degradation due to TCP's retransmission and congestion handling mechanisms. File transfer applications have higher bandwidth requirements but are not sensitive to delay and jitter - the user cares only about the end result. You may be familiar with the File Transfer Protocol (FTP) which relies on TCP, and the Trivial File Transfer Protocol, which uses UDP. Even so, TFTP has the same high bandwidth, low sensitivity as FTP. This traffic rarely requires special QoS treatment.

UDP /Voice / Video

Time-sensitive traffic like Voice and Video requires prioritization over other traffic so that it is transmitted immediately. This prioritization ensures that traffic transmitted at a constant rate by an endpoint such as an IP phone continues to be transmitted at a constant rate throughout the network, minimizing jitter.

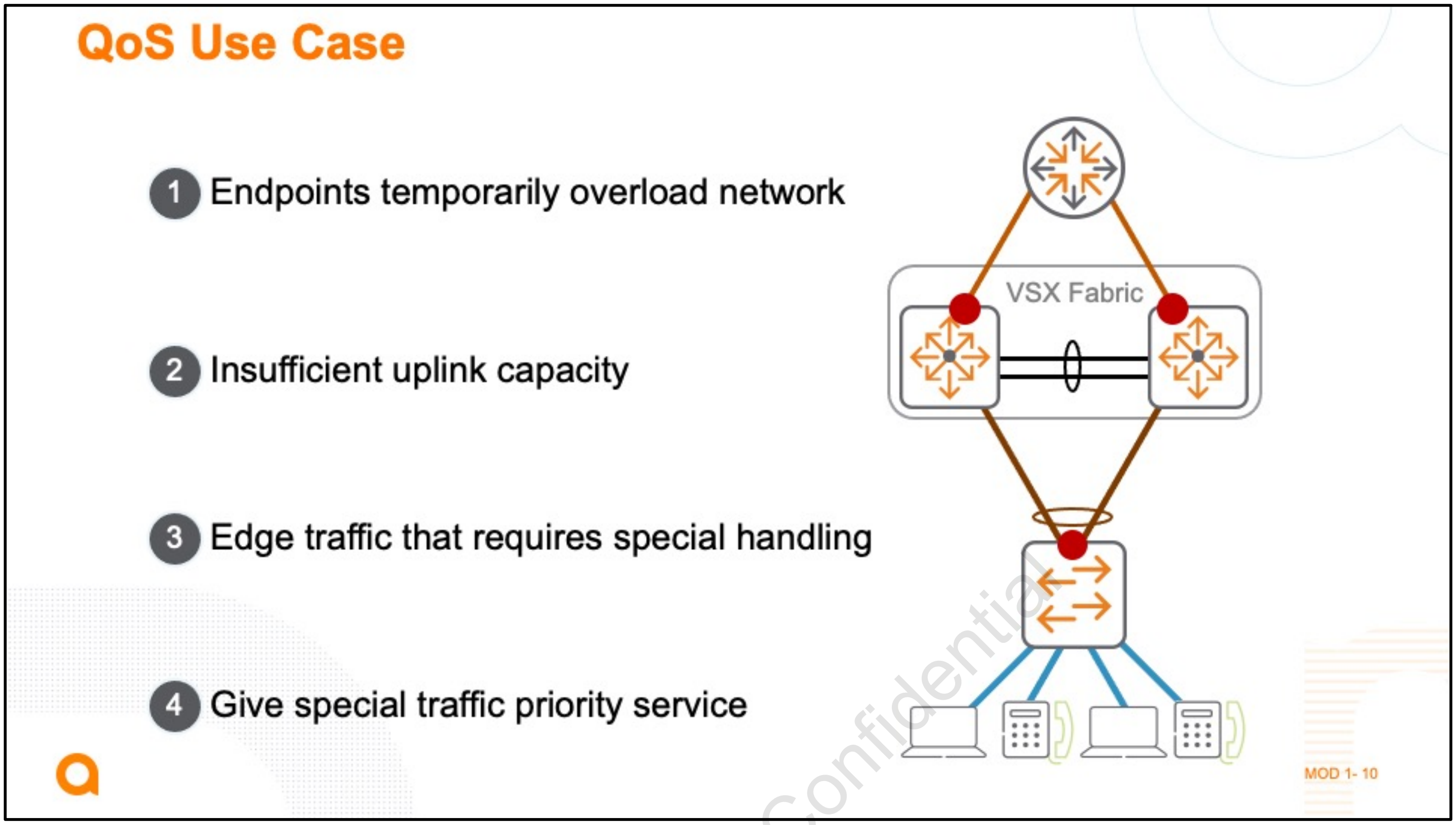
This traffic is nearly always a candidate for QoS prioritization. But remember, Less important UDP traffic must be prevented from monopolizing a link because the application will not back off on its own. You might want to enforce rate limits that ensure certain classes of traffic use only their fair share of bandwidth.

You know that UDP does not respond well to congestion, but often it doesn't matter. Suppose you transmit a video at 30 frames per second (fps) and one packet, containing one small part of one frame is dropped. It is unlikely anyone will notice – no need to retransmit video or voice traffic. Just move on without it. However, if too many packets are lost at once, user experience suffers. Users have a better experience when the network infrastructure responds to congestion gracefully by dropping less important or sensitive traffic.

Delay over media

Delay is the interval between the time that a message leaves the sender's mouth and the time that it reaches the receiver's ear. Some delay is inevitable, especially in transmissions over significant distances. While the speed of light in a vacuum is 186,000 miles per second (267,000 km per second), signals move through fiber or copper cable at approximately 125,000 miles per second (200,000 km per second). Consequently, a fiber network stretching halfway around the Earth introduces a delay of approximately 70 milliseconds under the best of circumstances with additional delay introduced by the physical transport medium. During periods of congestion, the delay can be much longer.

The effects of delay upon application performance depend heavily upon the expectations of users. The International Telecommunications Union (ITU) recommends no more than 150ms end-to-end delay for voice traffic. Greater delay can cause conversation collisions when each party begins to speak because the line appears quiet. Most users will find this situation to be unacceptable.



QoS Use Case

1. Endpoints temporarily overload network
2. Insufficient uplink capacity
3. Edge traffic that requires special handling
4. Give special traffic priority service

The figure shows times when you should implement QoS. Understand that use cases 1 – 4 are all a part of typical scenarios. Remember, well-designed networks can experience temporary bursts of congestion, during which there will be insufficient uplink capacity. During these times, some edge traffic may require special handling – you want to give special traffic some type of priority service.

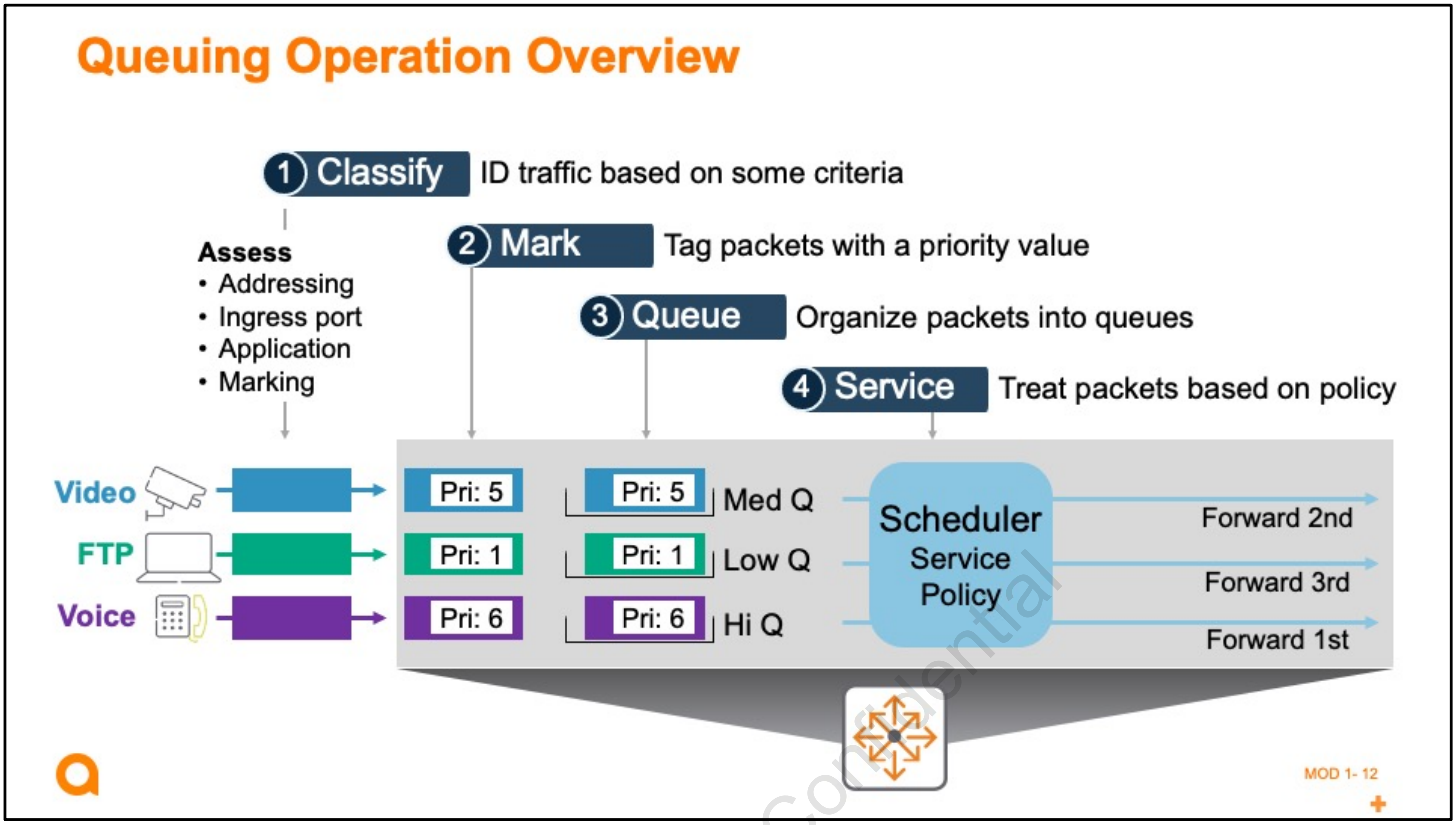
Understand that if your network experiences near-constant or very frequent congestion, QoS will not help you. You need faster links and/or more capable devices. QoS is for well-designed networks that experience relatively short, temporary bursts of congestion.

It is wise to think about QoS for any network deployment. Many network engineers think of this as a typical part of network design. During the almost inevitable short bursts of temporary over-subscription, which traffic in your network is business-critical? Which traffic requires high-bandwidth, low packet drop, and/or is sensitive to delay and jitter? When you can answer these questions, you are ready to then implement QoS – often by configuring network devices to classify traffic and apply QoS policies.

Classifying Traffic and Applying Policies

MOD 1- 11

Let’s examine traffic classification and policies.


Queuing Operation Overview
1 Classify ID traffic based on some criteria
2 Mark Tag packets with a priority value
3 Queue Organize packets into queues
4 Service Treat packets based on policy
Assess
• Addressing
• Ingress port
• Application
• Marking
Video
FTP
Voice
Pri: 5
Pri: 1
Pri: 6
Med Q
Low Q
Hi Q
Scheduler
Service Policy
Forward 2nd
Forward 3rd
Forward 1st
MOD 1- 12

Now you understand what QoS is, you should understand how it works – with a high-level, slightly simplified QoS scenario. Packets ingress at some switch or router interface – some video packets, some FTP packets, and some voice packets.

| First the switch must classify packets. It can assess source/destination MAC and IP addresses, the ingress port, the application type, and more, depending on the scenario. Perhaps all packets arriving on port 1/1/1 are low priority, or all VoIP packets are high priority.

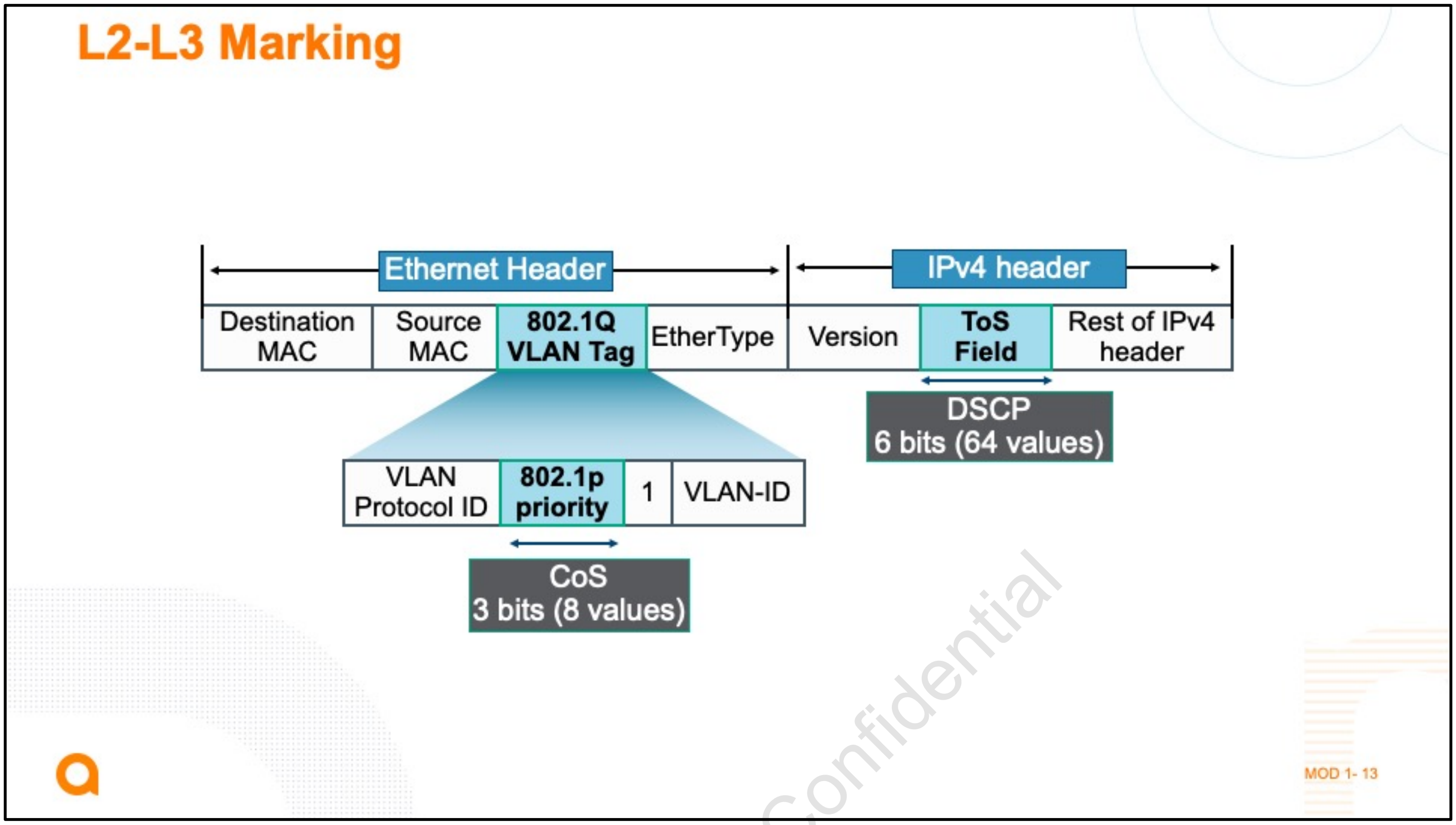
The purpose of classification is marking. Suppose you arrive at a music performance. An attendant classifies you based on your identification and your entry ticket. Then you are “marked” with a wrist band. Certain folks are on the performer’s guest list and are given a gold “all access” band. Others paid a premium for their ticket and are given a silver band – they get a special VIP line to quickly get food and beverages. Others paid a low price, are given a blue band, and must stand in a long line to get warm lemonade.

| In the figure, some packets were identified as video. These packets are marked upon ingress by placing a priority value of 5 in a special frame or packet header. Other packets classified as FTP and are marked with Priority=1. The VoIP traffic is marked with Priority=6. Notice that the classification process can also use marking. Perhaps an upstream device has already marked packets, or a VoIP phone has pre-marked its packets.

| Then packets are queued up for service. Video packets are placed in the medium queue, FTP in the low queue, and VoIP in the High Queue.

| Now the Scheduler takes over, often based on a configured Service Policy. Perhaps the policy states that the high queue shall be serviced exclusively, until it is empty. The router services other queues if and only if the high queue is empty. It goes back and forth between the medium and low queues, spending 70% more time servicing the medium queue.

Let’s look at classification and marking in more detail.



L2-L3 Marking

MOD 1- 13

To communicate the desired priority for traffic between network infrastructure devices, the devices use QoS marking. These mechanisms define specific values, which request specific priorities for traffic marked with that value. AOS-CX switches can leverage these values to determine how to classify and service traffic. The values reside within frames or packet headers and so are passed on with the traffic from device to device.

At Layer 2, Ethernet frames can have 802.1Q information, used for VLAN tagging. One of the fields inside the 802.1Q fields is used to do QoS tagging, as defined in the 802.1p standard. This is a 3-bit field, and so supports 8 priority values – 0 – 7. These fields are called the Class of Service (CoS) bits. If you use 802.1p user priority field that is a part of the 802.1Q standard, you are said to using CoS marking.

At Layer 3, the IP header has a 8-bit Type of Service (ToS) field. Initially, 5 bits were unused, and so 3 bits were used for ToS or “IP Precedence”. Then more of these 8 bits were defined in a standard. If using this standard, you are said to be using Differentiated Services Code Point (DSCP) marking.

Best effort service

This is the simplest service type. All traffic is treated equally in a first-come, first-served manner. If the traffic load is low in relation to the capacity of the network links, then there is no need for the administrative complexity and costs of maintaining a more complex end-to-end policy. If you intend to use this method, you must ensure that your network is over-designed – a so-called over-provisioned network where all link speeds are much higher than peak loads on the network.

802.1p or CoS

CoS uses a three-bit field in the 802.1Q VLAN tag to indicate priority. It defines eight values between 0 and 7. Note that because the VLAN tag carries the priority value, native VLAN traffic cannot contain a CoS value. Although this example shows an Ethernet frame that encapsulates an IP packet, an Ethernet frame with an 802.1p field can encapsulate any Layer 3 protocol an Ethernet supports.

ToS and IP precedence

ToS is rarely used today, but it is helpful to understand what ToS is if you hear it referenced.

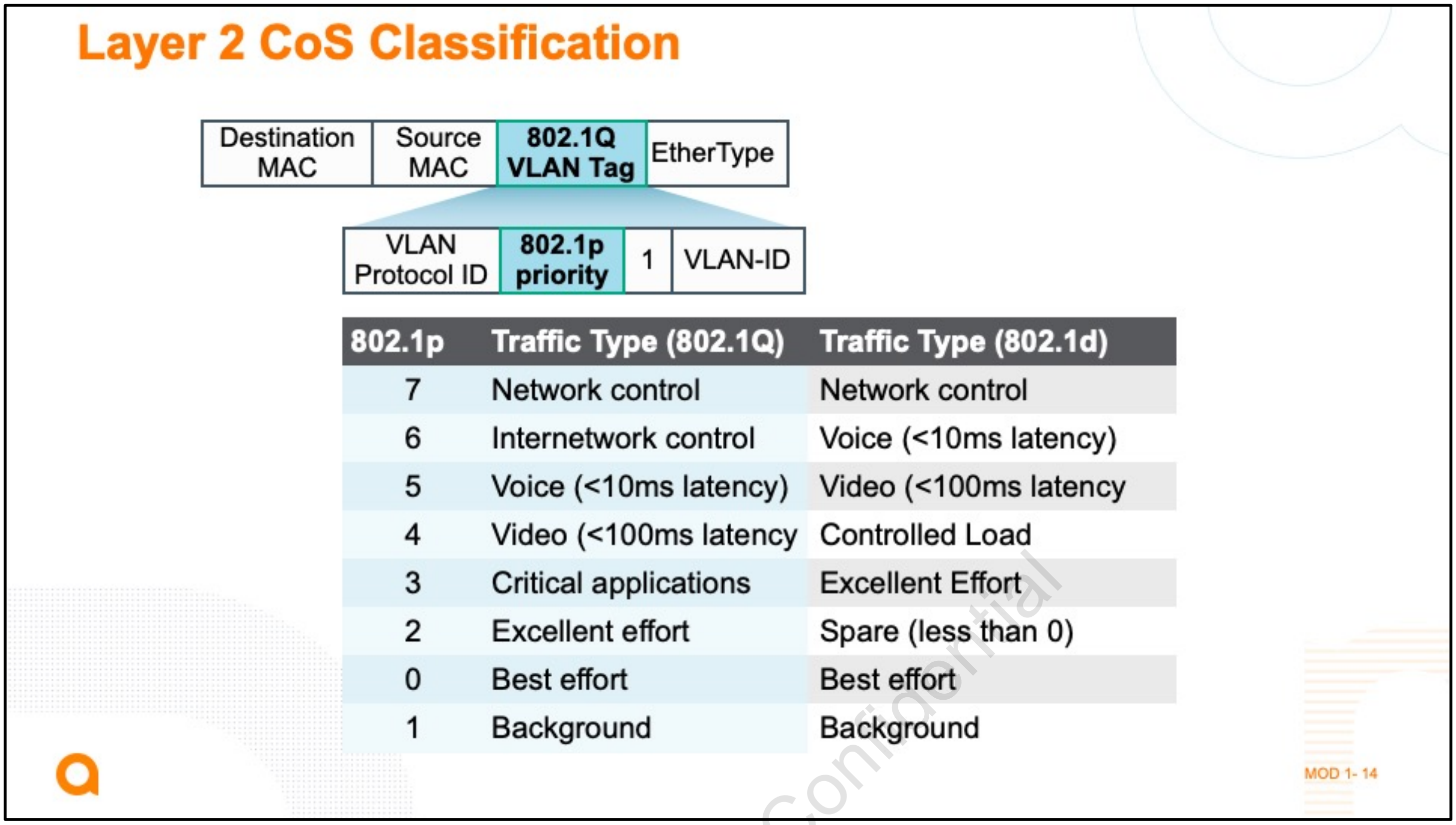
The figure shows that every IP packet includes an eight-bit field called the Type of Service (ToS) field. As originally defined, the ToS field contained two subfields—a three-bit IP precedence field used for relative prioritization and a four-bit subfield for the specific ToS desired by an individual packet. The remaining bit, bit 7, is called the must be zero (MBZ) bit and is unused.

The three IP precedence bits in the most significant position were to be used to provide eight levels of precedence or priority. The levels are labeled 0 to 7, with 7 being the highest priority and 0 the lowest. The values 6 and 7 are reserved for “Internet” and “network” uses, respectively. This enables routing updates and other critical traffic to have a higher priority than user-generated traffic. The highest user-defined value is 5, which represents “critical” priority.

The four ToS bits were designed to allow applications to instruct routers to choose routing paths with one of the following four characteristics: minimized delay, maximized throughput, maximized reliability, and minimized cost. Only one of the four bits could be “turned on,” that is, set to 1.

DiffServ / DSCP

Differentiated Services (DiffServ), defined in RFC 2474, supersedes a legacy IP QoS protocol called a type of Service (ToS). In DiffServ, the first six bits of the eight bit ToS field define 64 DiffServ “code points” (DSCPs). Instead of only defining priority relative to each other, the 64 DSCPs are also intended to define distinct forwarding behaviors, or Per Hop Behaviors (PHBs).



Layer 2 CoS Classification

Destination MAC	Source MAC	**802.1Q VLAN Tag**	EtherType

VLAN Protocol ID	**802.1p priority**	1	VLAN-ID

802.1p	Traffic Type (802.1Q)	Traffic Type (802.1d)
7	Network control	Network control
6	Internetwork control	Voice (<10ms latency)
5	Voice (<10ms latency)	Video (<100ms latency
4	Video (<100ms latency	Controlled Load
3	Critical applications	Excellent Effort
2	Excellent effort	Spare (less than 0)
0	Best effort	Best effort
1	Background	Background

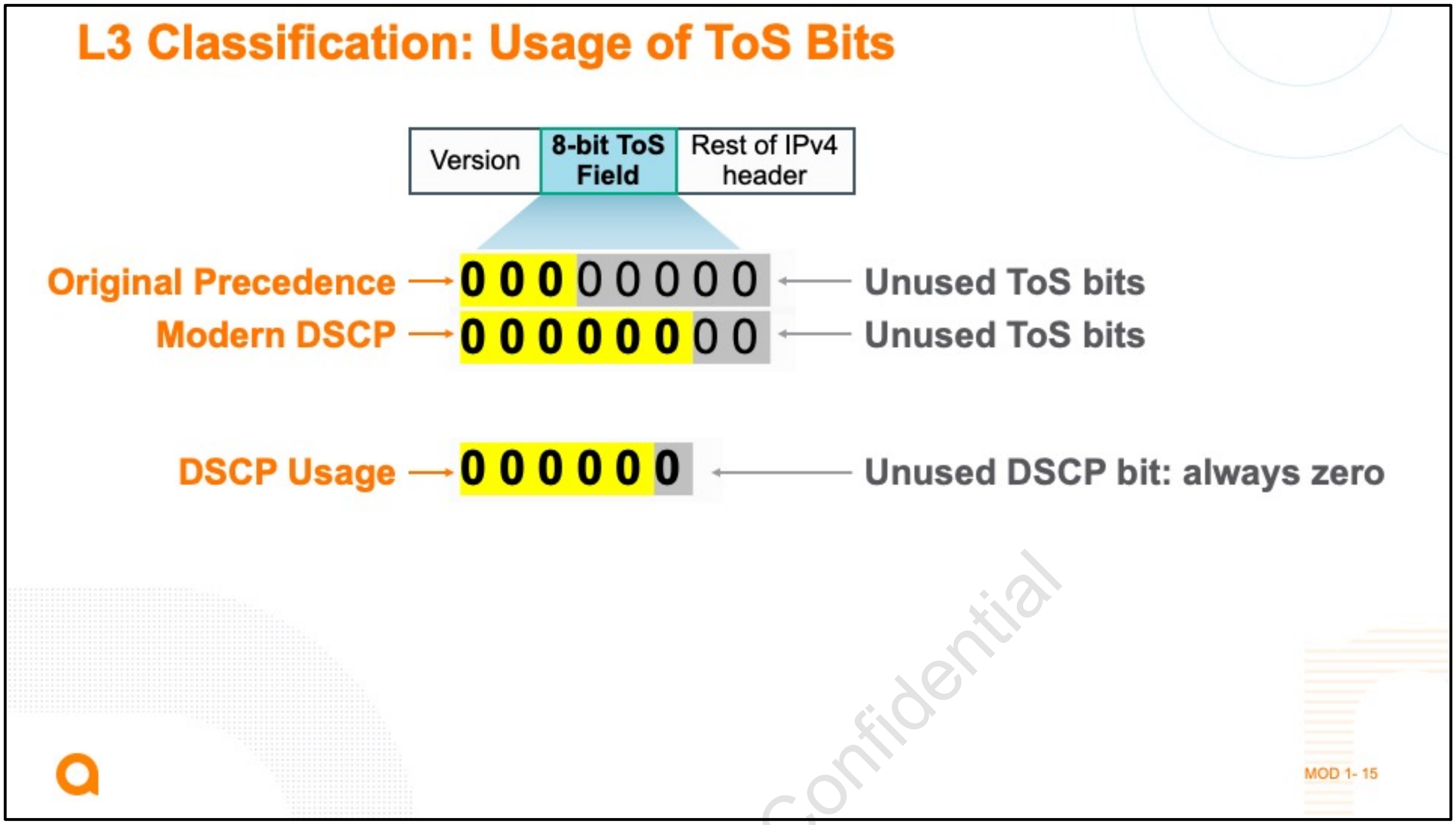
MOD 1- 14

CoS values request up to eight different priority levels for different types of traffic. Use the 3-bit Priority Code Point (PCP) 802.1p field within the 16-bit 802.1Q Ethernet VLAN tag to mark CoS. The standard recommendation differs based on whether the traffic type is 802.1Q or 802.1D.

For 802.1Q traffic, the you should use the most up-to-date recommendations in the IEEE 802.1Q-2014 standard. This standard recommends that you reserve the two highest values (7 and 6) for network control and internetwork control traffic. Network control traffic includes frames for layer 2 control protocols such as spanning tree BPDUs while internetwork control includes layer 3 control frames such as OSPF hellos and LSUs.

AOS-CX switches automatically mark their control traffic for high priority. Voice, value 5, and video, value 4, are given the next highest priority. These recommendations are based on their low latency requirements. Mission critical applications receive the next highest priority, value 3. The standard recommends assigning Value 2 to traffic that requires excellent effort. It is up to you to define what is considered critical or excellent effort applications. Next the switch gives its best effort to traffic with the default priority, which is 0. Finally, 802.1p value 1 is assigned to background traffic, which is less important or can tolerate delay. CoS 1 is deliberately set as the lowest CoS. This enables a traffic service level below the default (best effort) traffic level to be specified.

802.1D recommendations are based on the original standard that defines CoS. Some network infrastructures still use these recommendations, so you should be aware of them. In this scheme, network control traffic has just one value, 7 which shifts voice and video traffic up one priority level. Controlled load (rather than critical) has the next highest level and then excellent effort. But again, companies can choose themselves what applications fit in these categories. One of the most crucial differences in the 802.1D standard is that it assigns value 2 to "spare" traffic and is intended for traffic with a priority less than best effort. It is crucial that you understand whether other devices in the network consider value 2 as higher than 0 or less than 0.

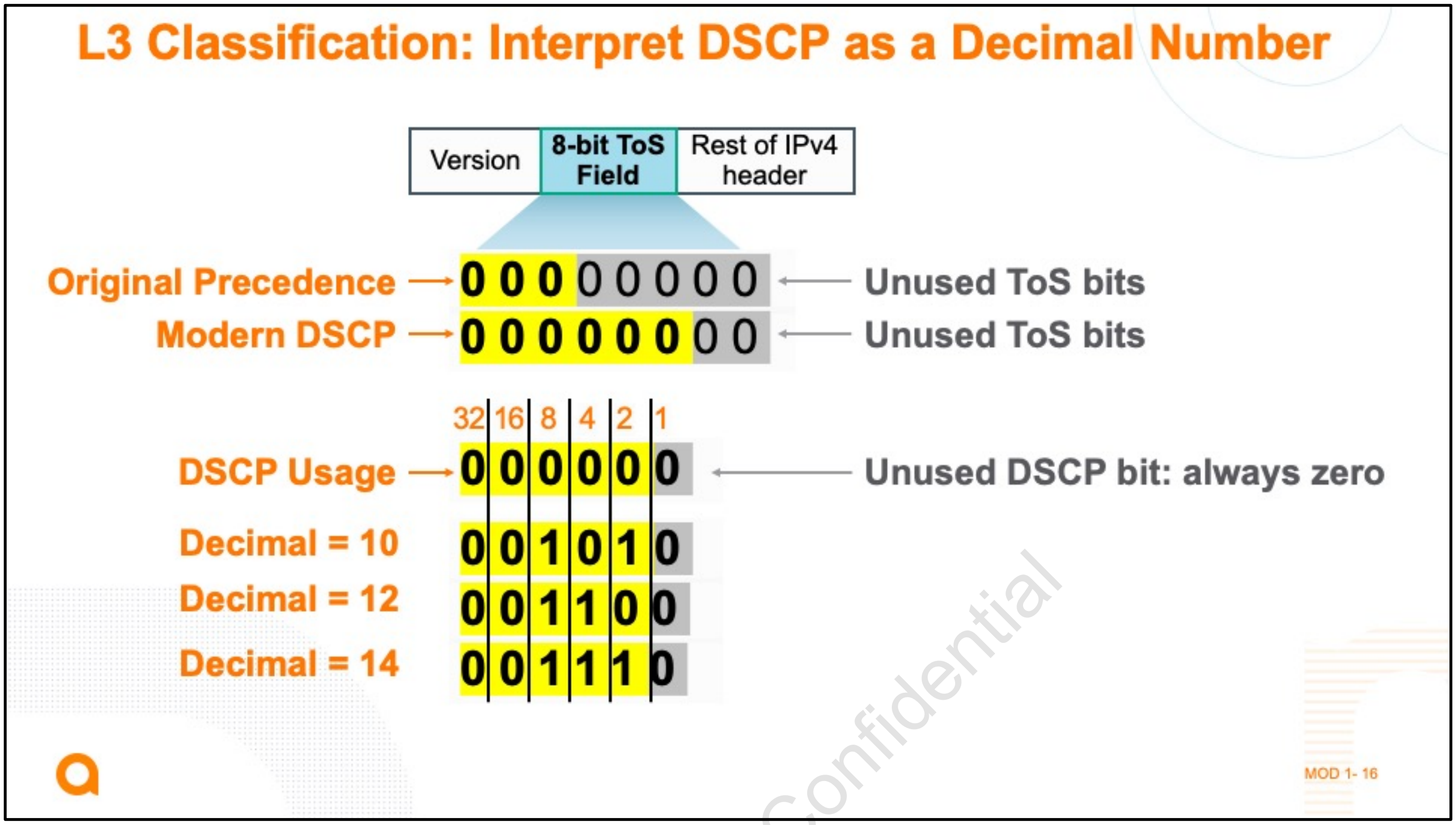


L3 Classification: Usage of ToS Bits

You know that every IP header includes an 8-bit ToS field. Originally, only the high-order 3 bits were used for IP Precedence. With three bits you have 8 possible values: decimal 0 – 7.

Much later, this 8-bit ToS field was redefined in a set of RFC standards to use 6 bits. Of those 6 bits, for nearly all common implementations, the lowest order bit is not used.

Note: Newer specifications can make specialized use of formerly unused bits, but they are not relevant to this discussion of QoS.



L3 Classification: Interpret DSCP as a Decimal Number

Version | 8-bit ToS Field | Rest of IPv4 header

Original Precedence → 0 0 0 0 0 0 0 0 ← Unused ToS bits

Modern DSCP → 0 0 0 0 0 0 0 0 ← Unused ToS bits

32 16 8 4 2 1

DSCP Usage → 0 0 0 0 0 0 ← Unused DSCP bit: always zero

Decimal = 10 0 0 1 0 1 0

Decimal = 12 0 0 1 1 0 0

Decimal = 14 0 0 1 1 1 0

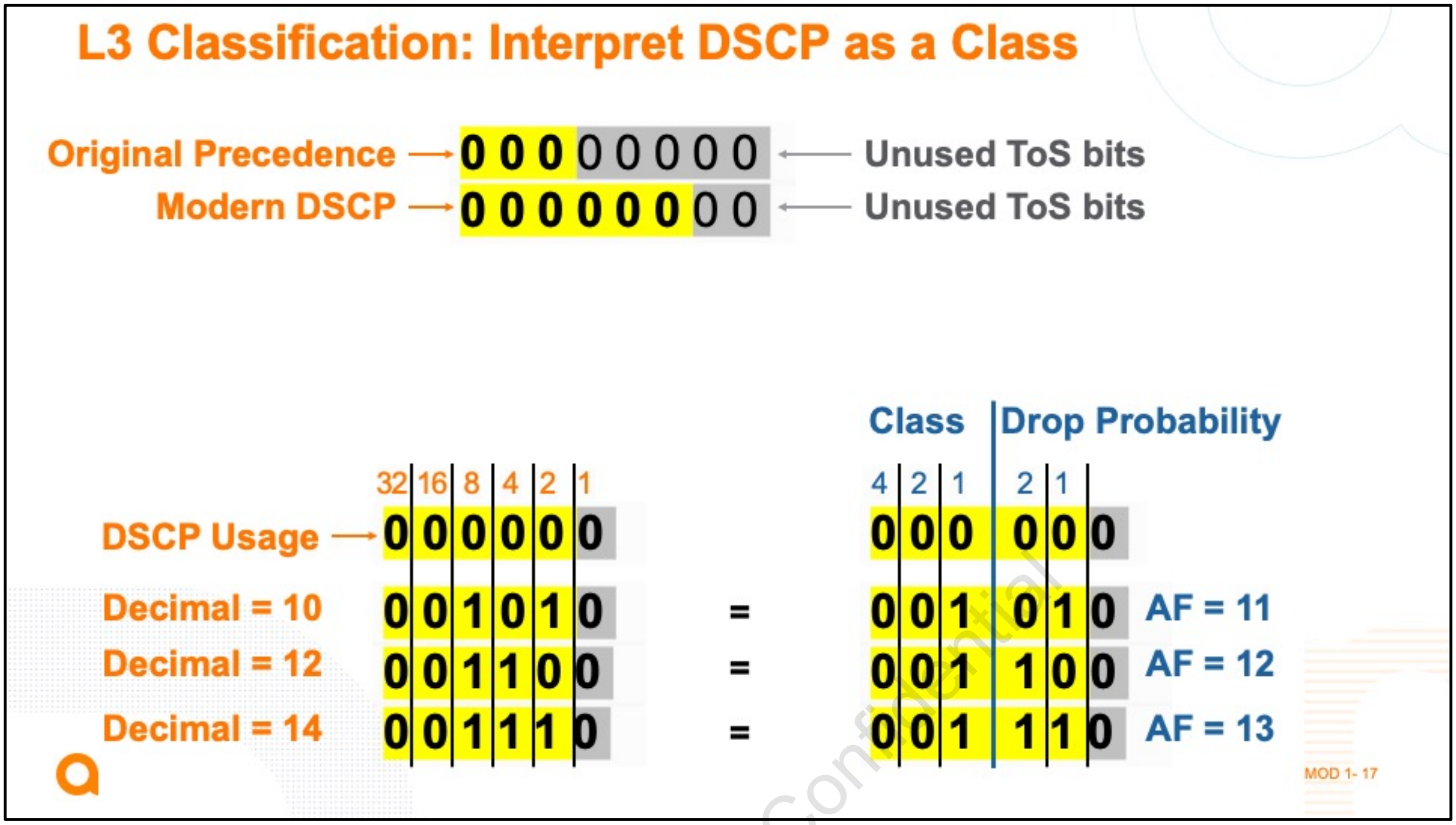
MOD 1- 16

Consider the decimal number 9,325. You know that the 5 is in the one's column, the 2 is in the 10's column, 3 is in the hundreds column, and 9 is in the thousands column.

With the 6-bit DSCP binary number, the columns are 1, 2, 4, 8, 16, and 32 – as shown in the figure. So, the binary number 001010 = 10 in decimal – there is a 1 in the eight's column, and a 1 in the 2's column, and 8 + 2 = 10.

Similarly, 001100 in binary = 12 in decimal, and 001110 = 14 in decimal.

So, when communicating with others you can say, "I have created a policy to mark certain packets with a DSCP decimal value of 10". (or 12, or 14, etc.) But there is a second way to interpret these bits.



L3 Classification: Interpret DSCP as a Class

Original Precedence → 0 0 0 0 0 0 0 0 ← Unused ToS bits

Modern DSCP → 0 0 0 0 0 0 0 0 ← Unused ToS bits

								Class			Drop Probability			
	32	16	8	4	2	1		4	2	1	2	1		
DSCP Usage →	0	0	0	0	0	0		0	0	0	0	0	0	
Decimal = 10	0	0	1	0	1	0	=	0	0	1	0	1	0	AF = 11
Decimal = 12	0	0	1	1	0	0	=	0	0	1	1	0	0	AF = 12
Decimal = 14	0	0	1	1	1	0	=	0	0	1	1	1	0	AF = 13

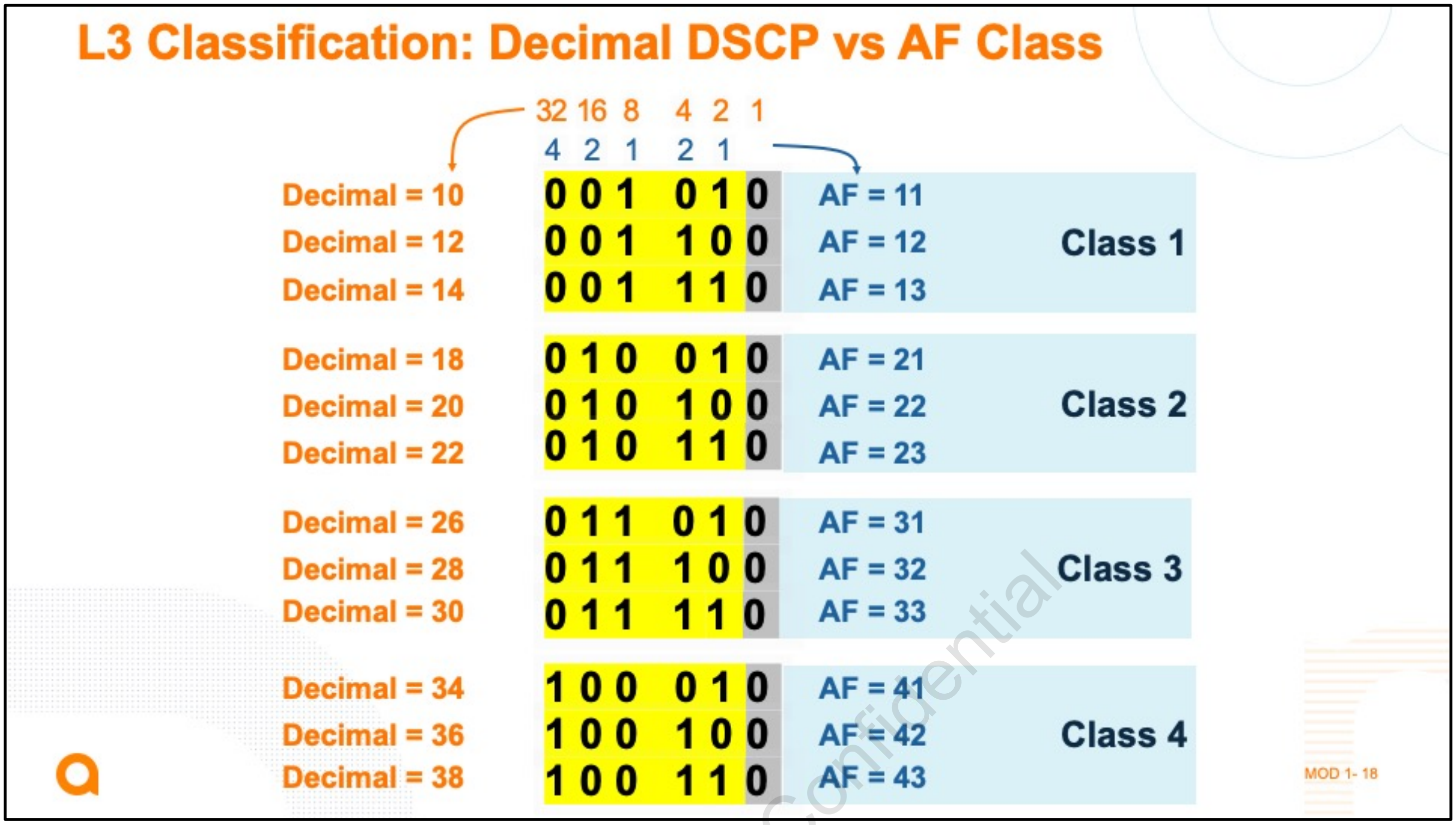
MOD 1- 17

Instead of perceiving the 6-bit number as a single value, think of the 6-bit field as two separate binary numbers. The three high-order bits represent the DSCP Class, and next two low-order bits represent a Drop Probability (DP), to further distinguish each class. That lowest order bit is simply ignored, as if it did not exist. Since they are two separate binary numbers, the high order bits have columns 1, 2, and 4, while the low-order bits have the 2 and 1 columns.

Now convert from binary to decimal as normal. Any packets marked with 001 in the high-order bits are in Assured Forwarding (AF) Class 1. If those packets also have 01 in the low-order bits, the packets are said to be in class AF 11. If the five used bits are set to 001 10, AF 12 is indicated. If the bits are 001 11, that is AF 13.

Packets with high-order bits 010 are in Assured Forwarding Class 2. Depending on the DP bits, they could be AF 21, AF22, or AF23.

Look at how this plays out.



L3 Classification: Decimal DSCP vs AF Class

	32 16 8	4 2 1		
	4 2 1	2 1		
Decimal = 10	0 0 1	0 1 0	AF = 11	Class 1
Decimal = 12	0 0 1	1 0 0	AF = 12	
Decimal = 14	0 0 1	1 1 0	AF = 13	
Decimal = 18	0 1 0	0 1 0	AF = 21	Class 2
Decimal = 20	0 1 0	1 0 0	AF = 22	
Decimal = 22	0 1 0	1 1 0	AF = 23	
Decimal = 26	0 1 1	0 1 0	AF = 31	Class 3
Decimal = 28	0 1 1	1 0 0	AF = 32	
Decimal = 30	0 1 1	1 1 0	AF = 33	
Decimal = 34	1 0 0	0 1 0	AF = 41	Class 4
Decimal = 36	1 0 0	1 0 0	AF = 42	
Decimal = 38	1 0 0	1 1 0	AF = 43	

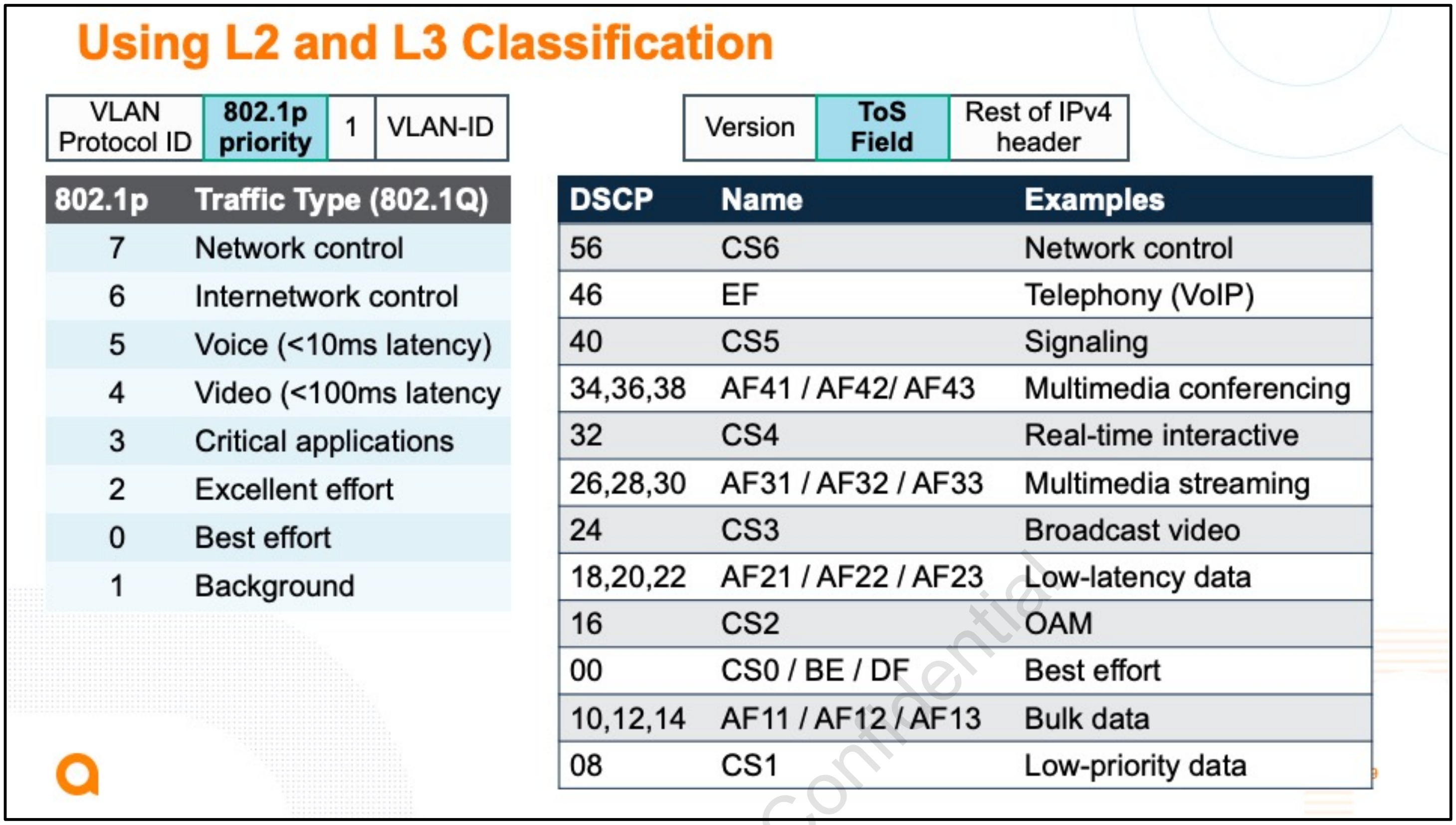
MOD 1- 18

Here are the four DSCP Assured Forwarding (AF) Classes, each with 3 Drop Probability values, as defined in RFC 2597. Each value is also interpreted as its equivalent decimal number. Spend some time looking at this. Patterns will become obvious to you, and it will become more intuitive.

Also, there is yet another way to interpret these bits, but its easy – just for backward compatibility with IP Precedence - the original, now largely unused practice of using only the three high order bits of the IP header's ToS field. Just focus on the highest order three bits in the figure, with columns 1, 2, and 4 – just like for the DSCP class, but without any lower order drop probability bits. You can simply count from 0 (binary 000) to 7 (binary 111). 000 is called DSCP "Class Selector (CS) 0", 001 is "CS1", and so on up to 111 – "CS7". DSCP and IP precedence can exist in parallel because of this interpretation.

If some applications only look for the three high order bits, they will recognize the IP Precedence value. An application or device that uses the DiffServ interpretation will interpret those same three bits from its own perspective. Applications and devices will arrive at the same relative priority, although the actual values are different.

You may be thinking, "Why are there two ways to talk about the same values that do the same thing?". Perhaps to accommodate differing preferences of different folks. The reason why is not for us to dwell upon. What is important is that we understand how these values are to be used in the real world.



Using L2 and L3 Classification

VLAN Protocol ID	**802.1p priority**	1	VLAN-ID

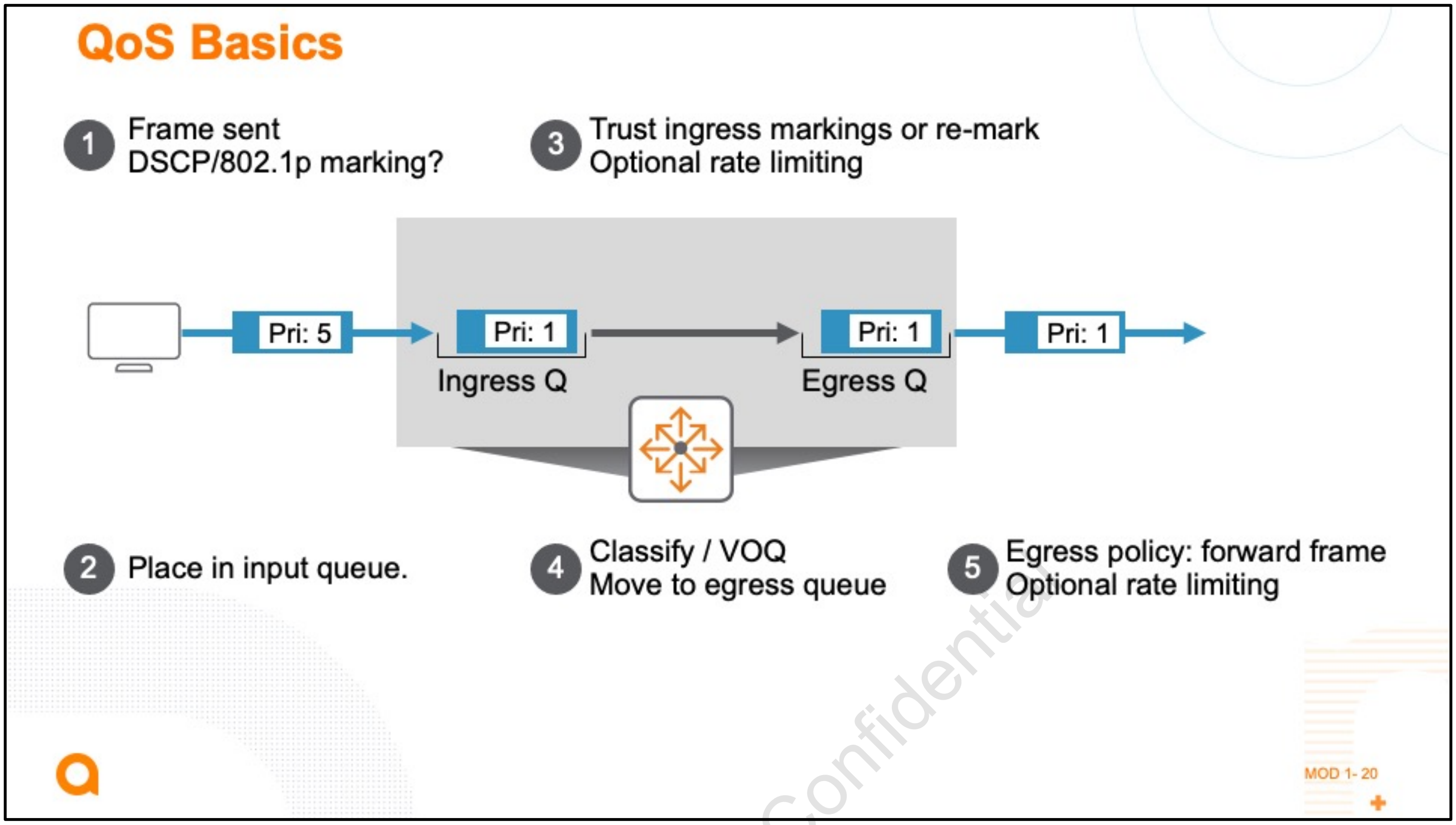
Version	**ToS Field**	Rest of IPv4 header

802.1p	Traffic Type (802.1Q)
7	Network control
6	Internetwork control
5	Voice (<10ms latency)
4	Video (<100ms latency
3	Critical applications
2	Excellent effort
0	Best effort
1	Background

DSCP	Name	Examples
56	CS6	Network control
46	EF	Telephony (VoIP)
40	CS5	Signaling
34,36,38	AF41 / AF42/ AF43	Multimedia conferencing
32	CS4	Real-time interactive
26,28,30	AF31 / AF32 / AF33	Multimedia streaming
24	CS3	Broadcast video
18,20,22	AF21 / AF22 / AF23	Low-latency data
16	CS2	OAM
00	CS0 / BE / DF	Best effort
10,12,14	AF11 / AF12 / AF13	Bulk data
08	CS1	Low-priority data

The table recommends the type of applications that you might assign to each AF class. For example, you could assign multimedia conferencing applications to AF class 4. These applications are interactive, so they need very low latency. You could assign multimedia streaming, which can use buffering and has a bit less strict latency requirements to AF class 3. You can assign AF class 2 to applications that require low latency, but less than multimedia. For example, interactive network management applications such as SSH could use this class. Finally, AF class 1 is for bulk data that can tolerate delays—it has priority less than 0.

If needed you can use the CS values for legacy IP precedence values. Also, RFC 3246 defines Expedited Forwarding (EF) class 46, to be used for applications most sensitive to jitter and delay. Aruba often recommends this DSCP for voice over IP (VoIP) traffic.



QoS Basics

1 Frame sent
DSCP/802.1p marking?

3 Trust ingress markings or re-mark
Optional rate limiting

Pri: 5

Pri: 1
Ingress Q

Pri: 1
Egress Q

Pri: 1

2 Place in input queue.

4 Classify / VOQ
Move to egress queue

5 Egress policy: forward frame
Optional rate limiting

MOD 1- 20

This figure illustrates the basic steps that traffic experiences when you implement a QoS solution.

| A device sends a frame, with or without classification markings.

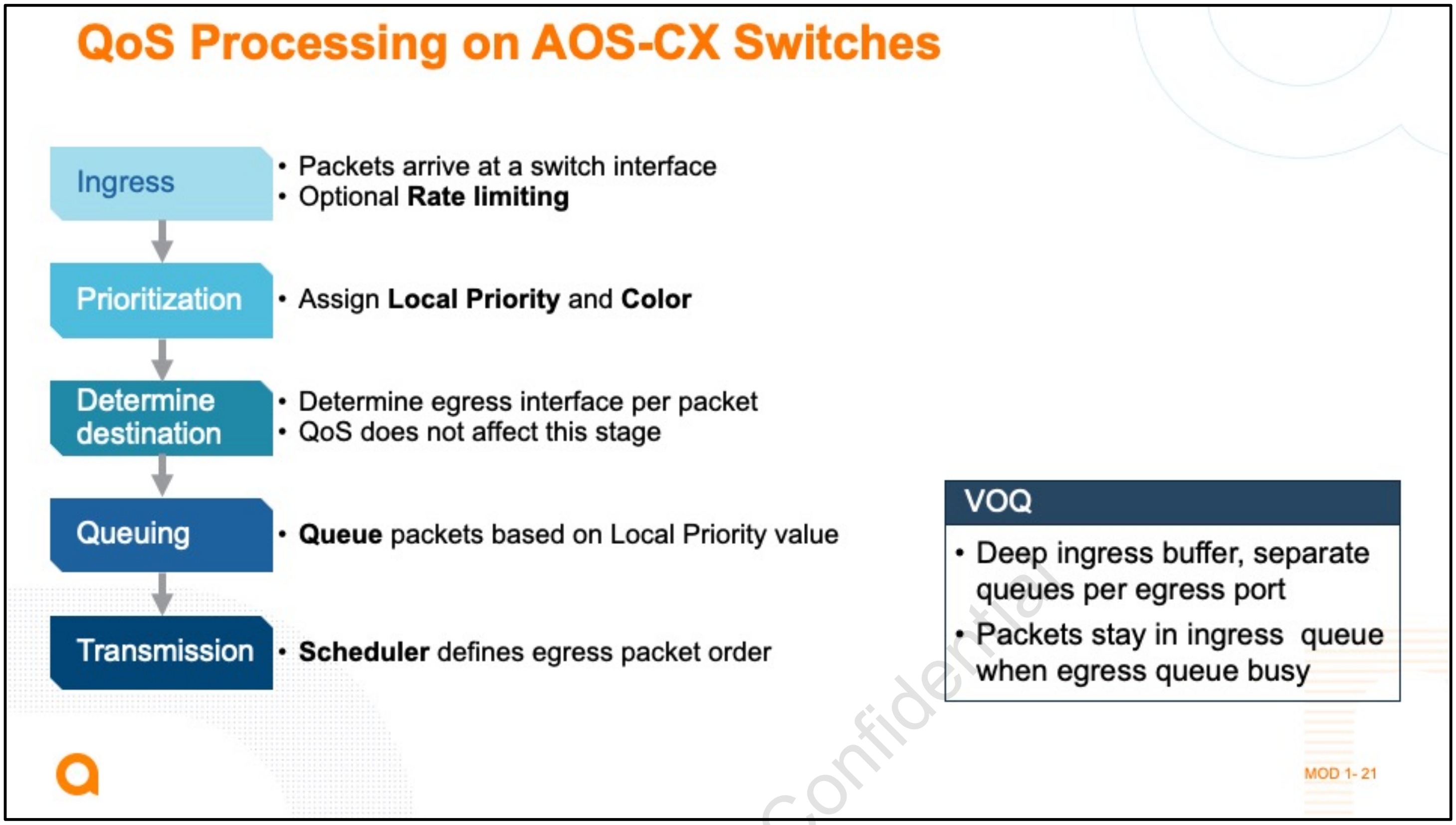
| The switch receives the frame and places it in the ingress queue.

The switch determines if the ingress packet was pre-marked by the host application. And if so, does the switch trust these markings. If there are not trusted markings, the switch can mark or re-mark the frame based on some policy you define. If ingress marking is trusted, the switch can use those markings for any policy you may have defined. If you expect both L3 DSCP and L2 802.1p markings, you should define a policy to control which marking to trust. You might also define an ingress rate limiting policy - applied certain ports.

| The switch then uses the classification information and Virtual Output Queuing (VOQ) to determine when to move the frame to the egress queue.

| The switch then implements the egress queuing policy, based on the traffic classification, to determine when the frame is sent out the port, along with applying any defined outbound rate limiter.

Let’s look at this process in a bit more detail.



QoS Processing on AOS-CX Switches

Ingress
- Packets arrive at a switch interface
- Optional **Rate limiting**

Prioritization
- Assign **Local Priority** and **Color**

Determine destination
- Determine egress interface per packet
- QoS does not affect this stage

Queuing
- **Queue** packets based on Local Priority value

Transmission
- **Scheduler** defines egress packet order

VOQ
- Deep ingress buffer, separate queues per egress port
- Packets stay in ingress queue when egress queue busy

MOD 1- 21

ArubaOS-CX switches use a special process to locally apply QoS features. This process has five stages:

Ingress. Packets that arrive at a switch interface could be filtered by the rate limit QoS feature. This QoS feature can limit Broadcast, Multicast and Unknown Unicast (BUM) traffic, to protect the switch from unnecessary processing.

Prioritization. The switch applies the QoS classifier tool to determine the correct priority for a specific traffic type. ArubaOS-CX defines two values: Local Priority and Color. The Local Priority helps to determine the correct egress priority queue for the traffic. The color determines the traffic's drop eligibility. If queues become congested, more drop-eligible traffic will be dropped first.

Destination Determination. The egress interface for each packet is determined based on the switching and routing logic. Internally the switch moves the traffic to the correct egress interface. QoS tools in this stage are not used and therefore do not affect this stage.

Queuing. The switch assigns each egress packet to an interface queue, based on the packet's Local Priority value. ArubaOS-CX supports up to 8 queues. Packets wait in the queue for their turn to be transmitted.

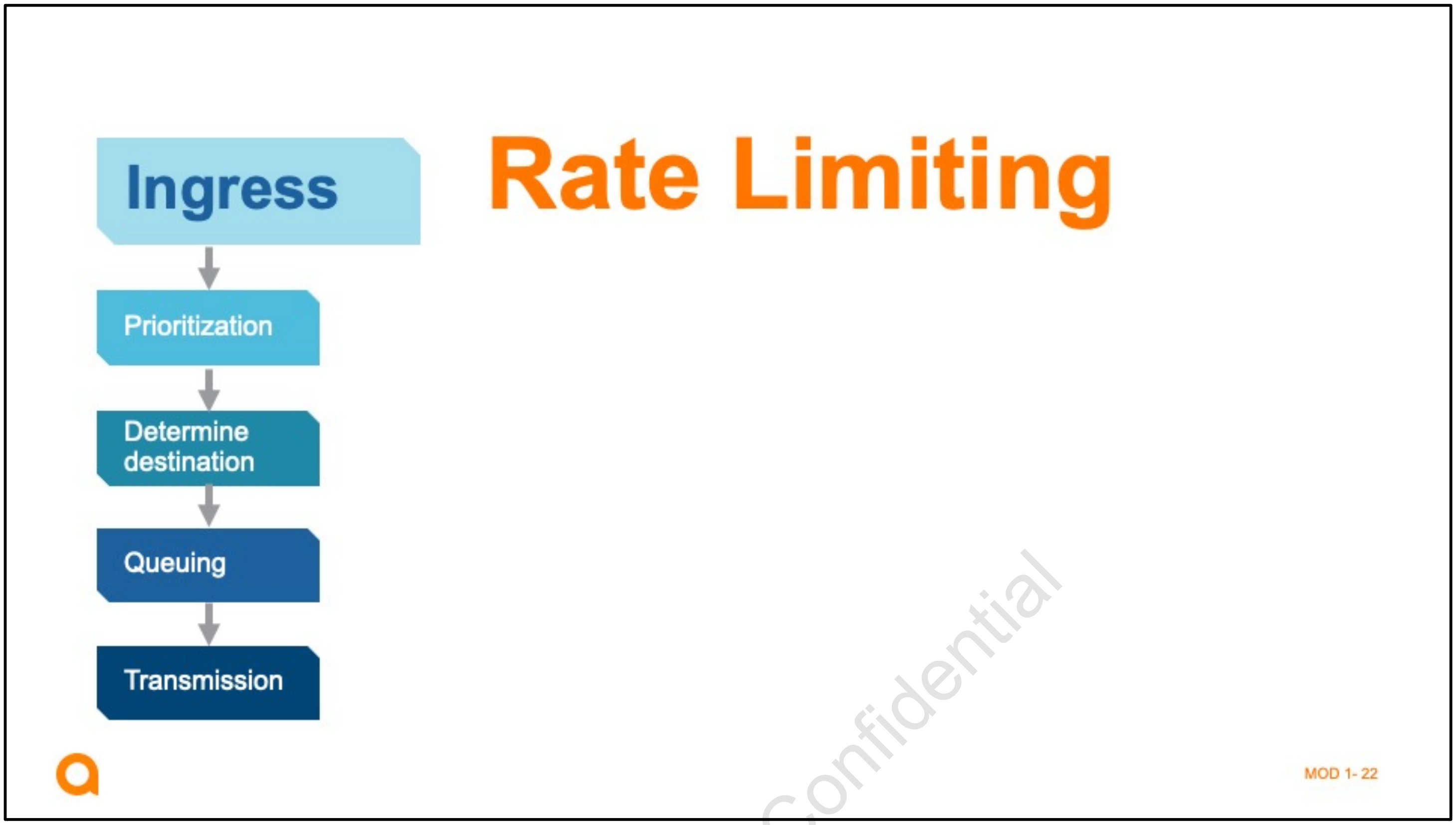
Transmission. In this stage a Scheduler QoS tool defines the order in which packets from different queues are transmitted out the egress interface.

ArubaOS-CX switches use an intra-switch queuing method called Virtual Output Queuing

(VoQ). If the ingress buffer used a single queue, head of line (HOL) blocking could delay the traffic. If the packet at the front of the queue is destined out a congested port, it delays all packets behind it, even though those that are destined to non-congested ports.

VoQ prevents this problem by providing deep ingress buffers with separate queues for each egress port. Physically the ingress buffer consists of internal and external DRAM banks with most traffic typically being buffered in the internal DRAM. This means that packets will stay in the inbound port queue when the outbound queue is under load. Using this technique packets are never drop in the outbound interface but in the inbound. This feature is embedded in the Operating System and is not configurable.

Let’s dive into each one of these processes.



Rate Limiting

Ingress

Prioritization

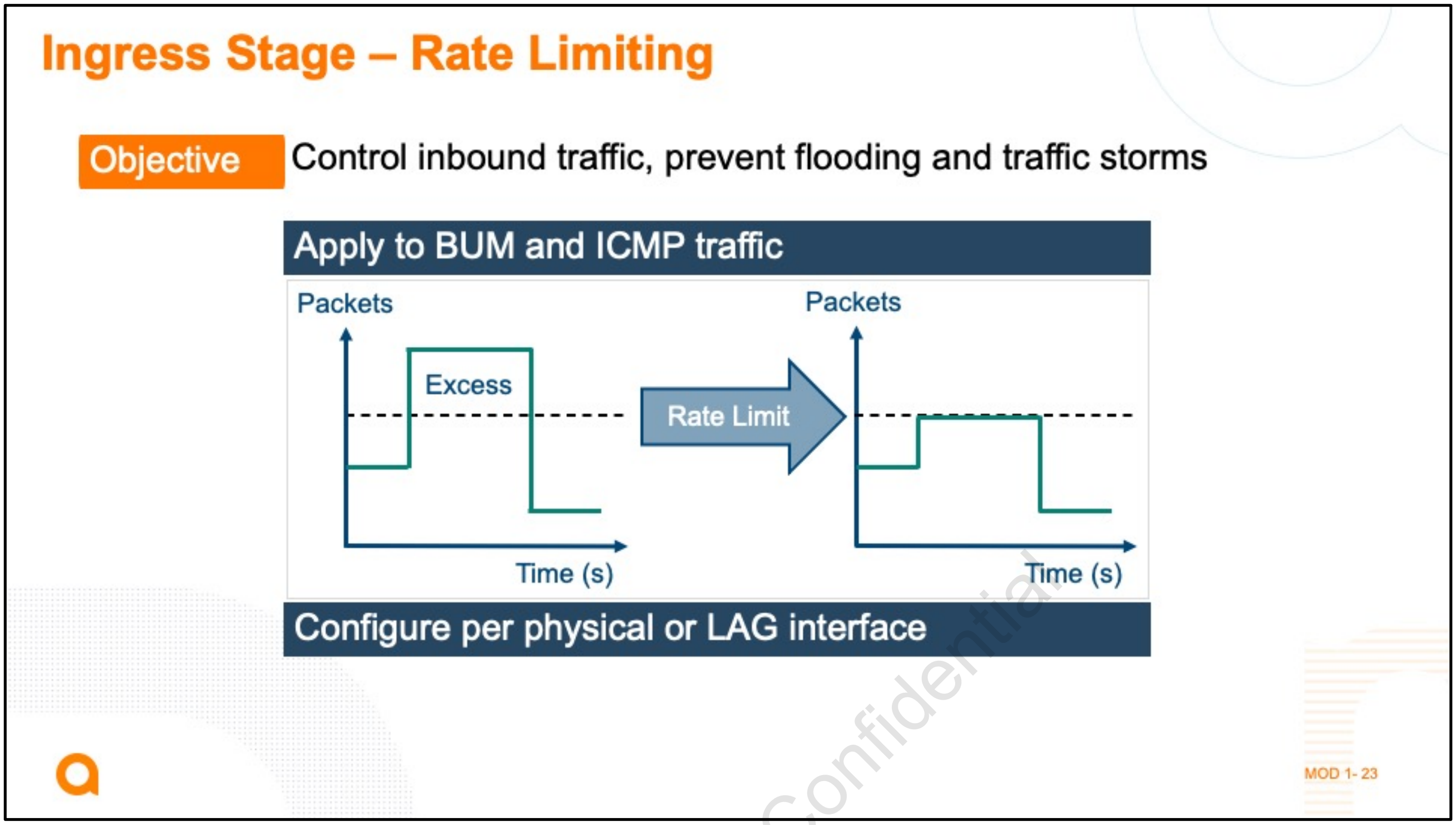
Determine destination

Queuing

Transmission

MOD 1- 22

Let’s focus on the first stage of the QoS process – Ingress Rate Limiting.



Ingress Stage – Rate Limiting

Objective Control inbound traffic, prevent flooding and traffic storms

Apply to BUM and ICMP traffic

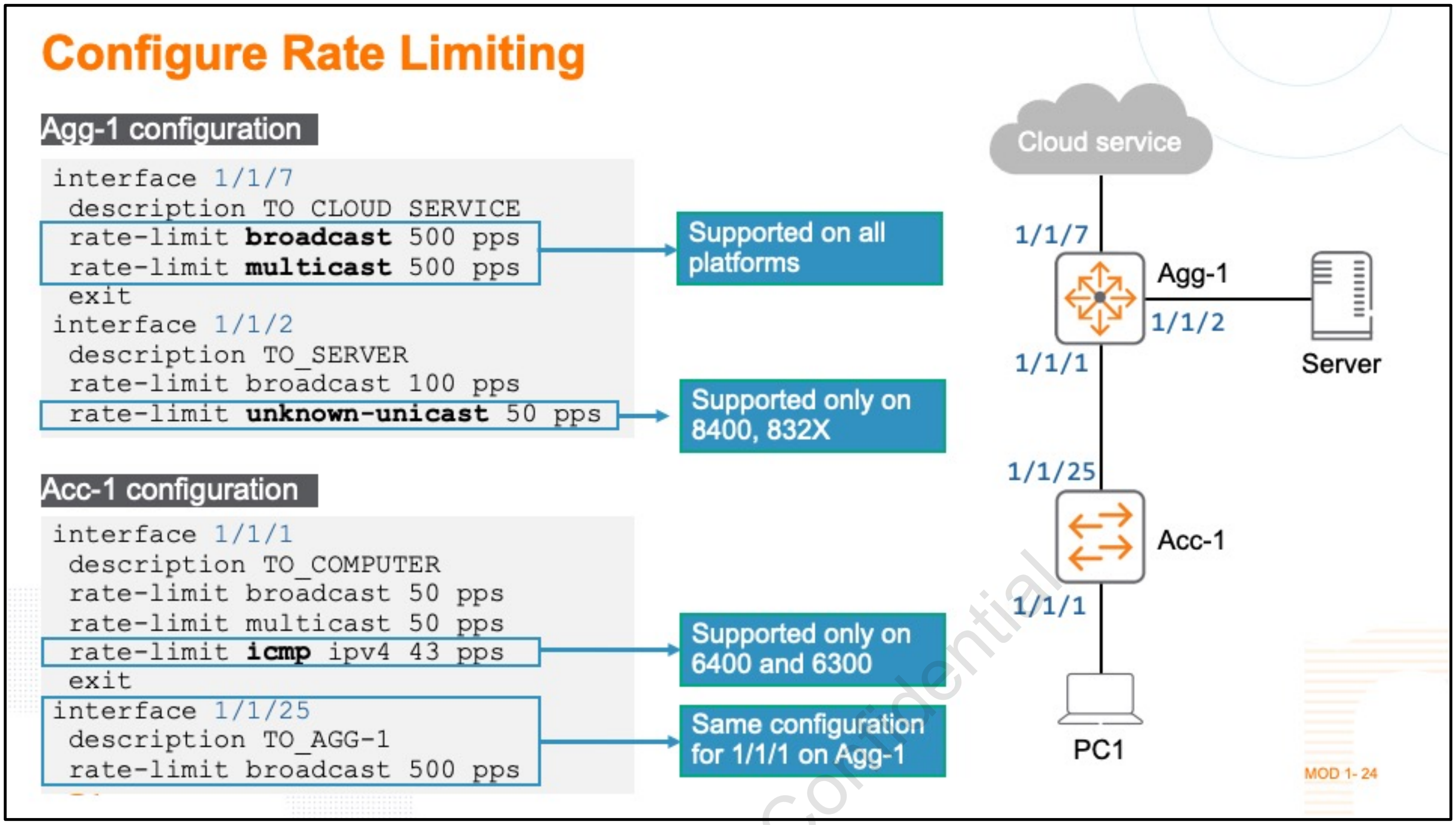
Configure per physical or LAG interface

MOD 1- 23

Port rate limiting helps control undesirable traffic. The goal is to allow only enough Broadcast, Unknown unicast, and Multicast (BUM) traffic for the network to function properly, while preventing flooding and traffic storms.

A certain amount of BUM traffic is required for normal network operation. Broadcast packets may include ARP and DHCP traffic, for instance. Video streams, and certain types of network protocol packets, are multicasts. Unknown-unicast packets may be intended for devices whose addresses have temporarily aged out of network-forwarding caches. Configuring rate limits can help provide the balance between necessary traffic and flooded traffic.

Rate Limiting is a traffic policer QoS tool. It ensures that traffic conforms to a defined rate – a bandwidth limit. An important characteristic of rate limiting is that it does not add delay to traffic to be forwarded - that which is under the threshold and is accepted by the switch. You configure it per physical or LAG interface.



Configure Rate Limiting

Agg-1 configuration

```
interface 1/1/7
 description TO CLOUD SERVICE
 rate-limit broadcast 500 pps
 rate-limit multicast 500 pps
 exit
interface 1/1/2
 description TO_SERVER
 rate-limit broadcast 100 pps
 rate-limit unknown-unicast 50 pps
```

Supported on all platforms

Supported only on 8400, 832X

Acc-1 configuration

```
interface 1/1/1
 description TO_COMPUTER
 rate-limit broadcast 50 pps
 rate-limit multicast 50 pps
 rate-limit icmp ipv4 43 pps
 exit
interface 1/1/25
 description TO_AGG-1
 rate-limit broadcast 500 pps
```

Supported only on 6400 and 6300

Same configuration for 1/1/1 on Agg-1

You can implement the rate limiting QoS feature on all ArubaOS-CX platforms. However, supported parameters are platform dependent:

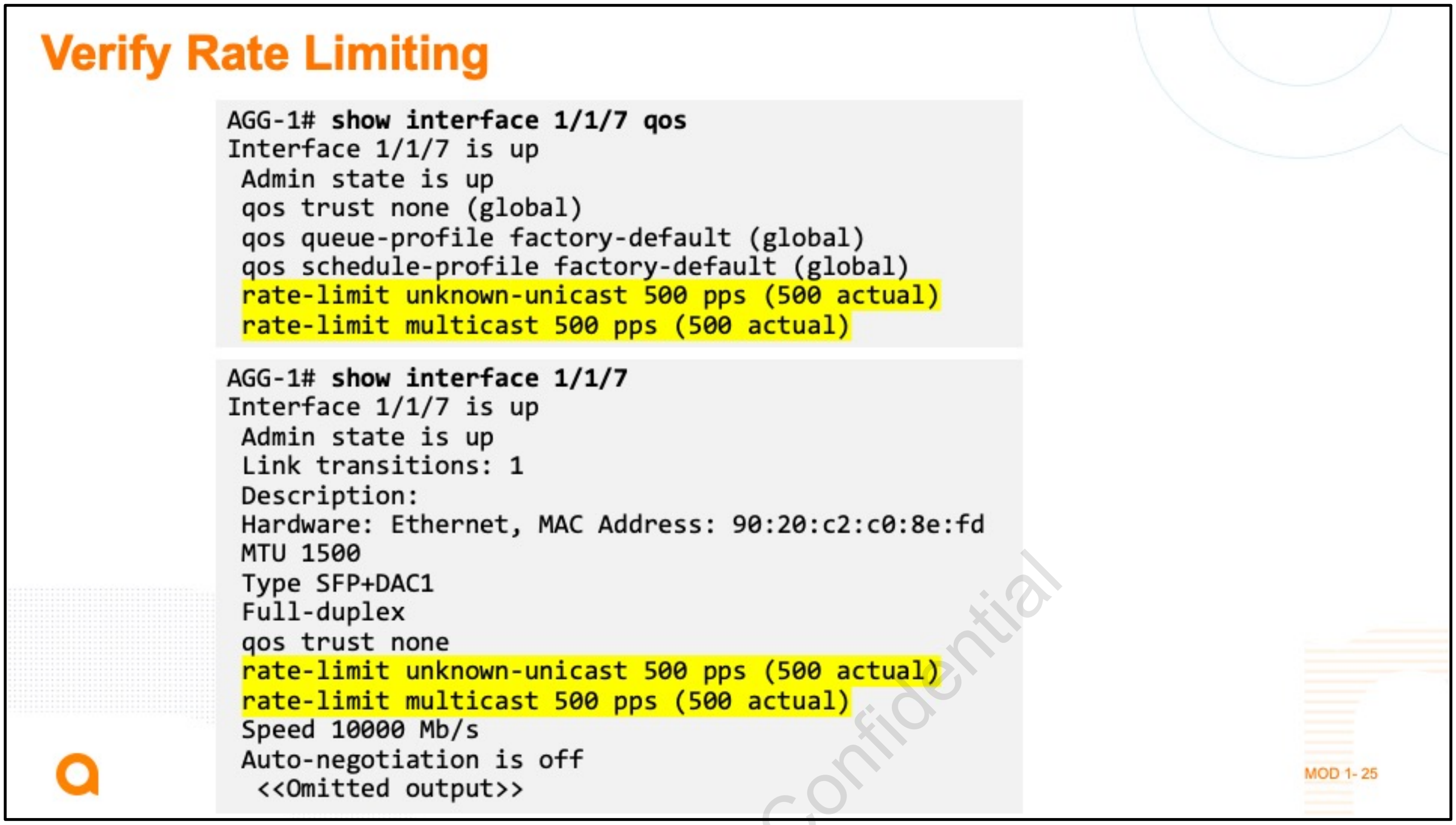
Broadcast: Supported on all ArubaOS-CX platforms, limit only the broadcast traffic.

Multicast: Supported on all ArubaOS-CX platforms, limit only the multicast traffic.

Unknown Unicast: Supported only on 8400, 8325 and 8320 switch platforms, limit only the unknown unicast traffic.

ICMP: Supported only on 6400 and 6300 switch platforms, the ICMP rate limit can be configured to apply to IPv4, IPv6, or all IP traffic. Only one ICMP rate-limit can be configured at a time. Applying a new ICMP rate-limit replaces any previous ICMP rate-limit.

Note: When multicast and broadcast rate limit are applied at the same time on the same interface, broadcast packets are limited to the lower of the two rate values.



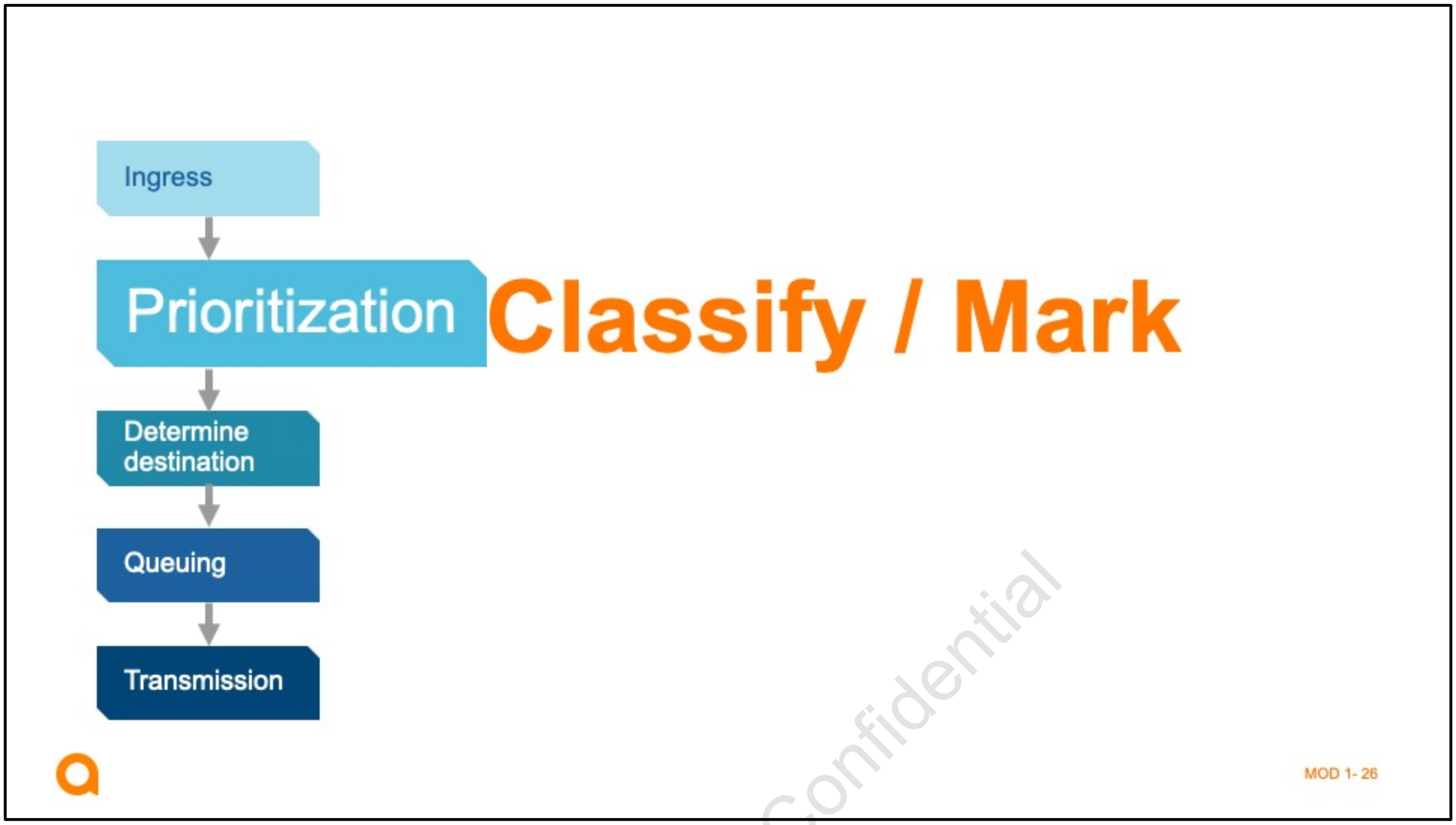
Verify Rate Limiting

```
AGG-1# show interface 1/1/7 qos
Interface 1/1/7 is up
 Admin state is up
 qos trust none (global)
 qos queue-profile factory-default (global)
 qos schedule-profile factory-default (global)
 rate-limit unknown-unicast 500 pps (500 actual)
 rate-limit multicast 500 pps (500 actual)
```

```
AGG-1# show interface 1/1/7
Interface 1/1/7 is up
 Admin state is up
 Link transitions: 1
 Description:
 Hardware: Ethernet, MAC Address: 90:20:c2:c0:8e:fd
 MTU 1500
 Type SFP+DAC1
 Full-duplex
 qos trust none
 rate-limit unknown-unicast 500 pps (500 actual)
 rate-limit multicast 500 pps (500 actual)
 Speed 10000 Mb/s
 Auto-negotiation is off
  <<Omitted output>>
```

MOD 1- 25

The show interface <interface-id> qos and show interface <interface-id> comands are useful to verify the rate-limiting feature implemented on the interface.



Ingress

Prioritization **Classify / Mark**

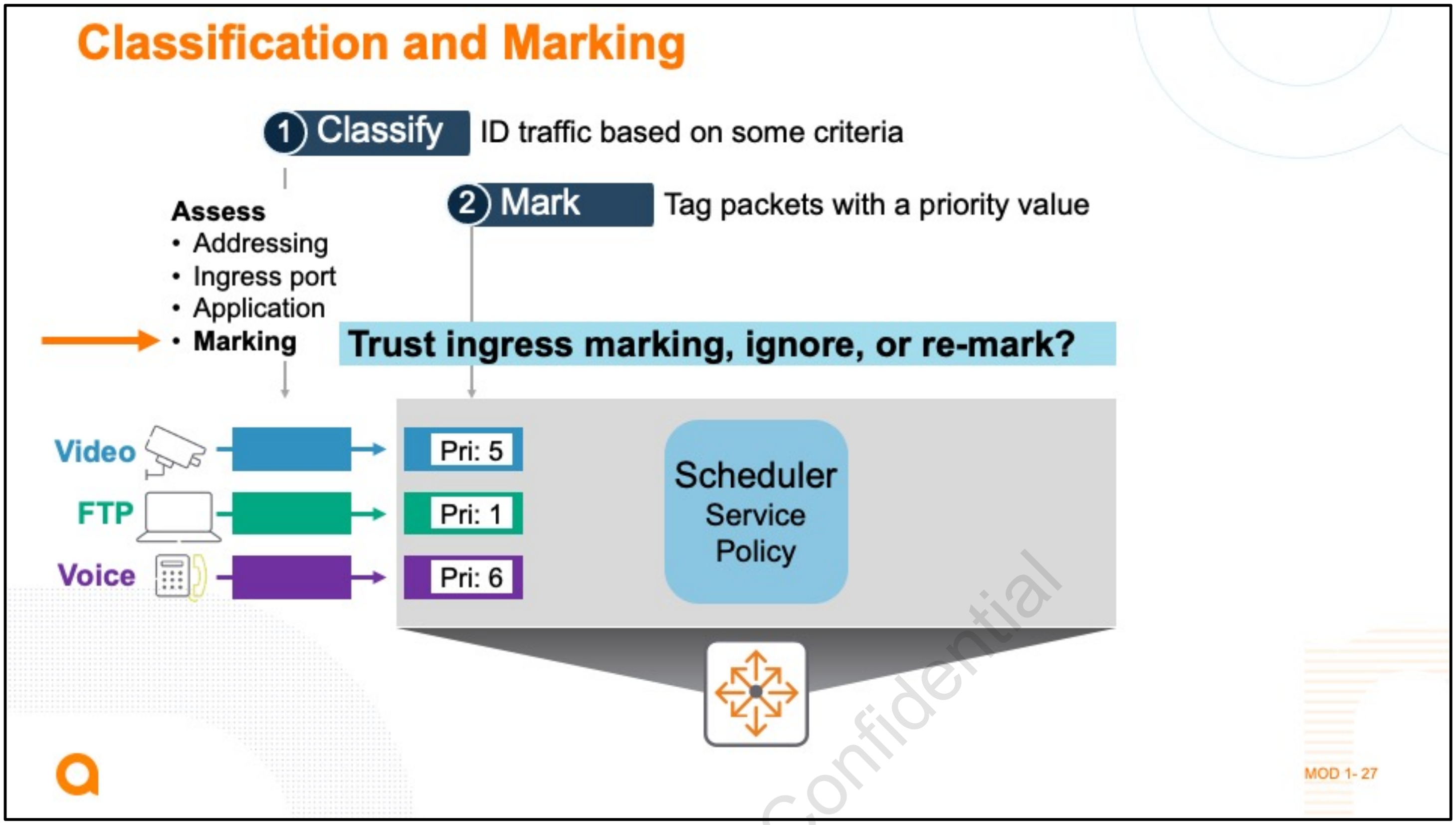
Determine destination

Queuing

Transmission

MOD 1- 26

Now we move to stage two – classification and marking.



Classification and Marking

1 Classify ID traffic based on some criteria

2 Mark Tag packets with a priority value

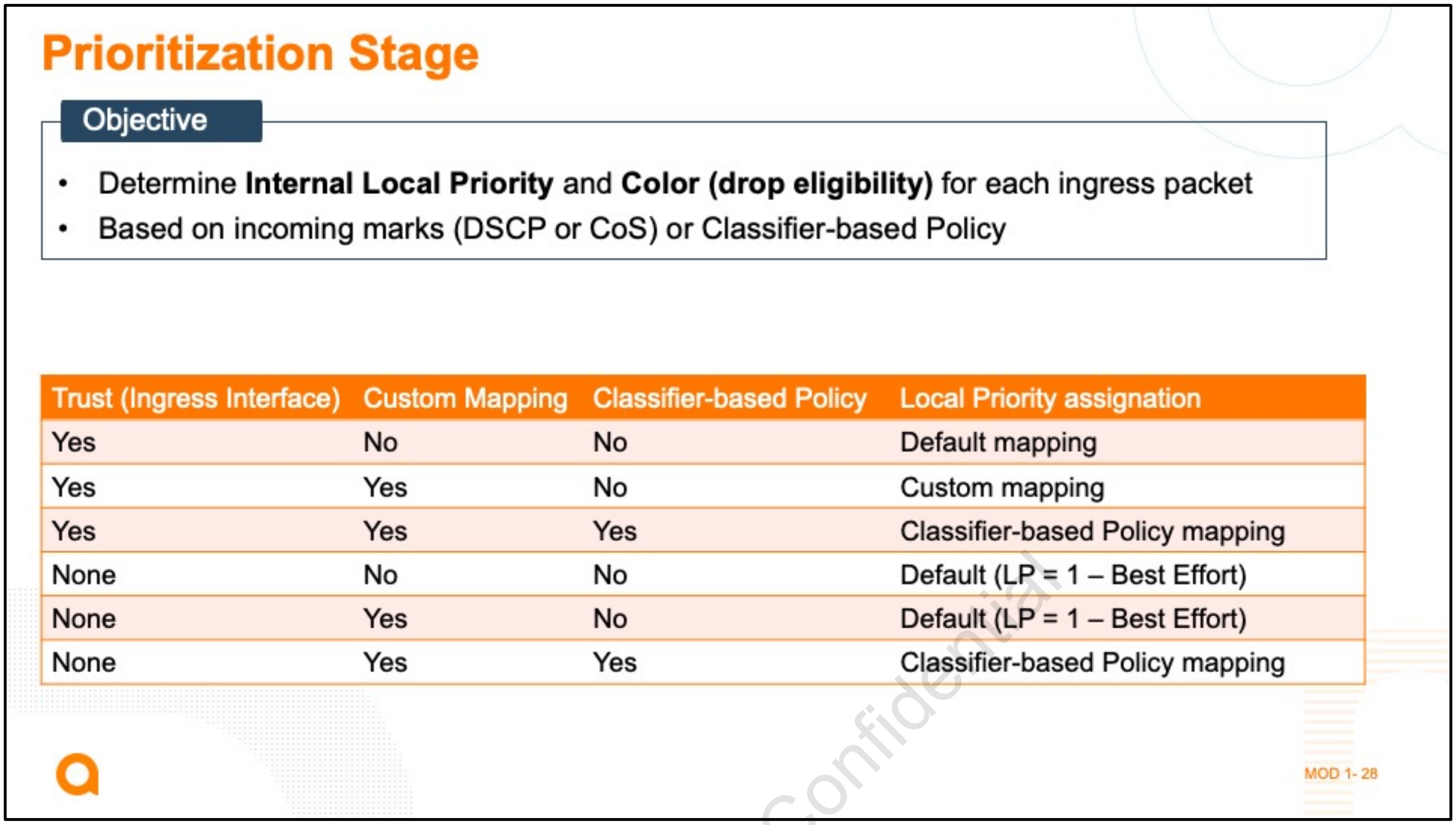
Assess
- Addressing
- Ingress port
- Application
- **Marking**

Trust ingress marking, ignore, or re-mark?

MOD 1- 27

Recall that classifiers perform a kind of identification process. They inspect incoming traffic, identify it, and then assign it to a class. You define a classifier-based policy, analyzes packet parameters to apply a specific class of service.

Notice that the classification process can also use marking. Perhaps an upstream device has already marked packets, or a VoIP phone has pre-marked its packets. You can configure certain switch ports to trust these markings, and take some action based on them – provide priority service levels perhaps. You can configure other ports to ignore ingress frame markings, and re-mark them according to a local policy that you define.



Prioritization Stage

Objective

- Determine **Internal Local Priority** and **Color (drop eligibility)** for each ingress packet
- Based on incoming marks (DSCP or CoS) or Classifier-based Policy

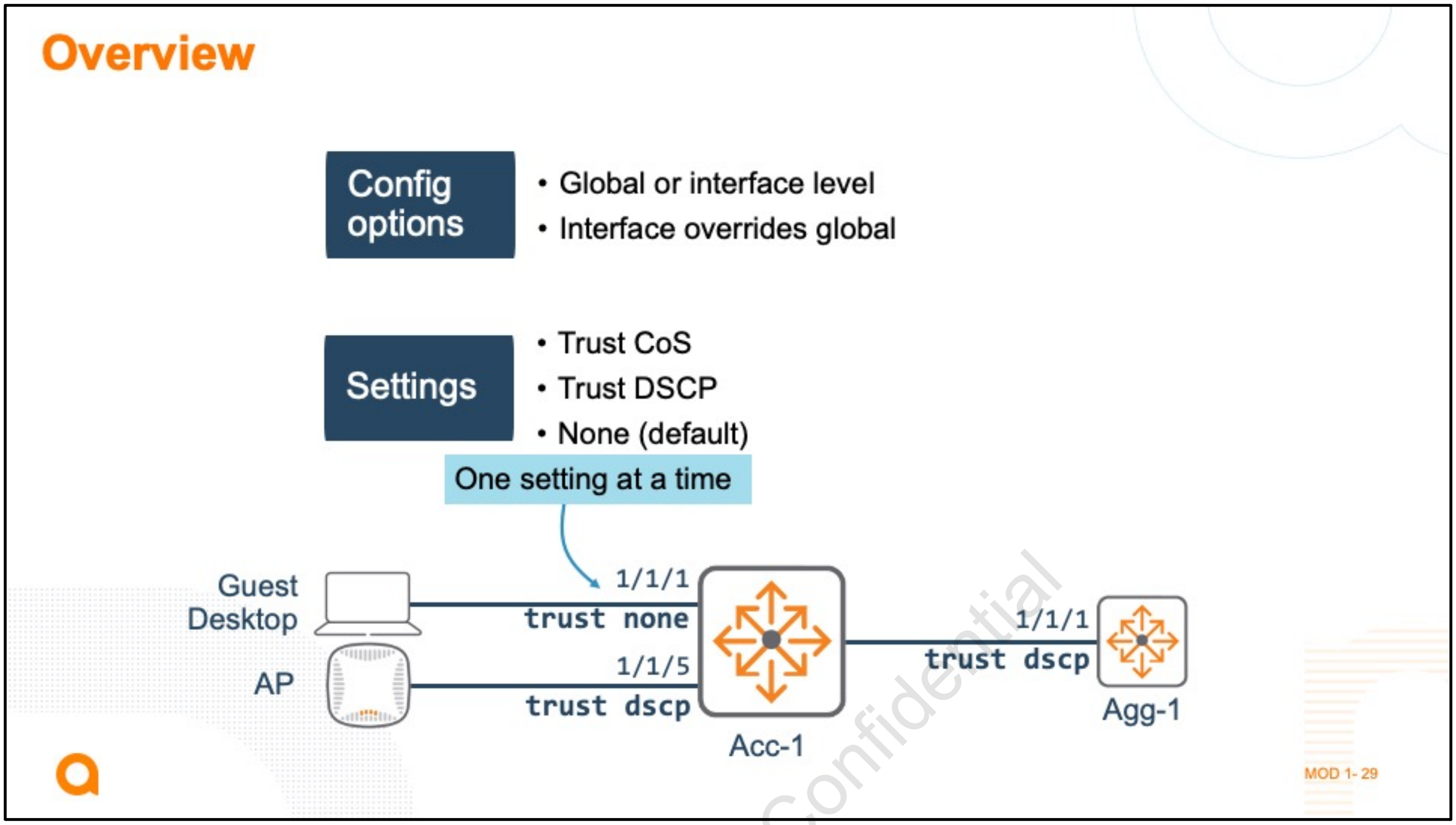
Trust (Ingress Interface)	Custom Mapping	Classifier-based Policy	Local Priority assignation
Yes	No	No	Default mapping
Yes	Yes	No	Custom mapping
Yes	Yes	Yes	Classifier-based Policy mapping
None	No	No	Default (LP = 1 – Best Effort)
None	Yes	No	Default (LP = 1 – Best Effort)
None	Yes	Yes	Classifier-based Policy mapping

MOD 1- 28

The objective of the prioritization stage is to determine the Internal Local priority value and the Color (drop eligibility) for a specific traffic. The ArubaOS-CX switch uses these parameters to specify the behavior, or the class of service used.

The decision to assign a Local Priority and Color value could be based on ingress packet L3 DSCP or L2 CoS marks. Each switch maintains a mapping table that helps with the decision. You can display the default mapping table with the show qos dscp-map default command or show qos cos-map default command. ArubaOS-CX supports the creation of custom mapping. The other possibility is to use a classifier-based policy where the interesting traffic is selected based on specific characteristics. In this case the marking may or may not be present.

Trust is an important setting for QoS in ArubaOS-CX. When it is set on the incoming interface the switch can use the incoming marks to determine the Local Priority. However, if the trust is not set (or set to none) then incoming marking cannot be used and all traffic is placed in the Default Best Effort behavior, represented by the Local Priority 1. Notice that classify policy-based assignment is not dependent on the trust status and overrides default and custom mapping. The table in the slide summaries this discussion.



Overview

Config options
- Global or interface level
- Interface overrides global

Settings
- Trust CoS
- Trust DSCP
- None (default)

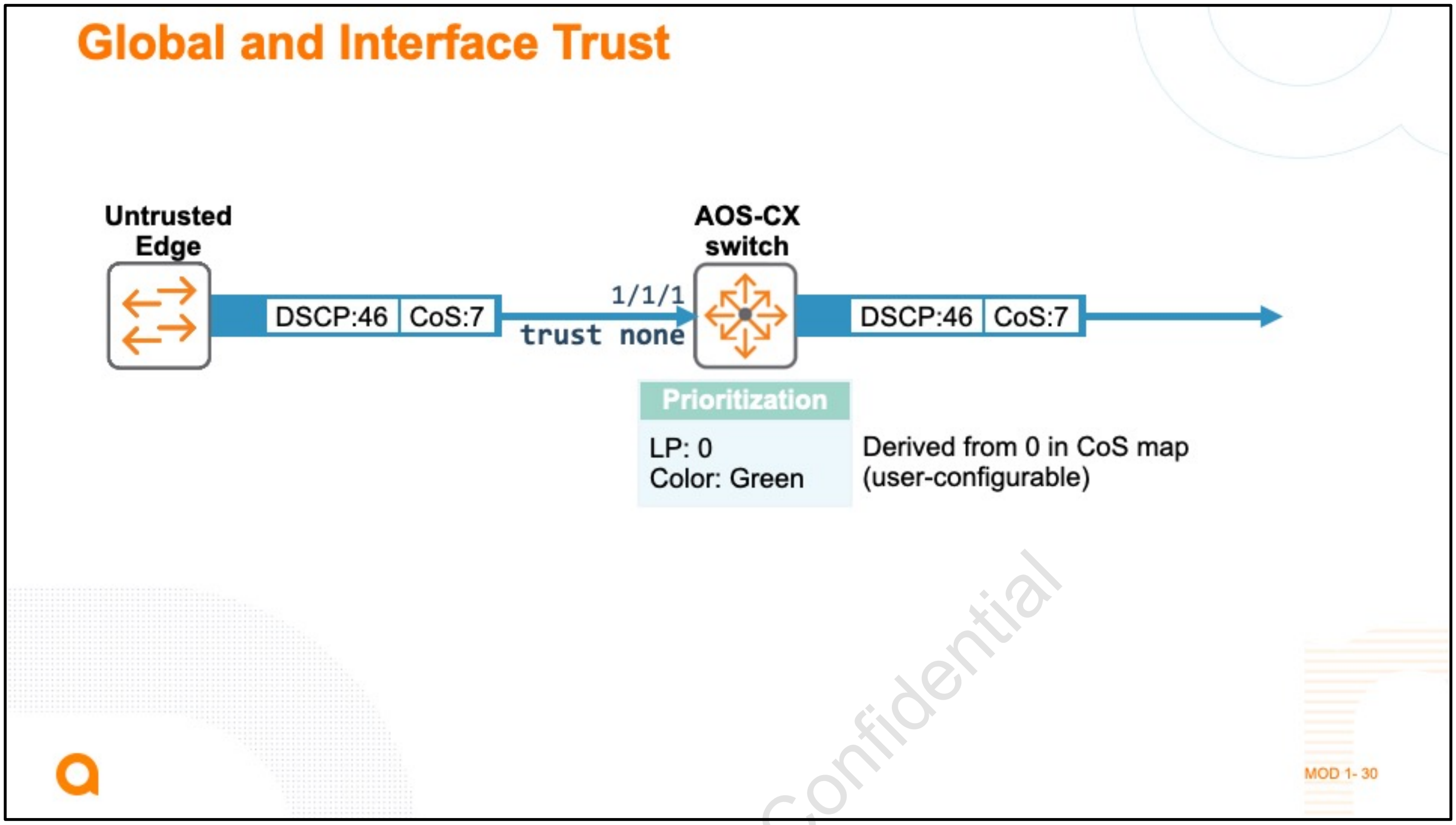
One setting at a time

When traffic arrives at the switch, the switch can trust existing markings or not. The term trust refers to whether the packet markings are valid input for the classifier, which decides each packet's service class.

You know that trust configuration plays an important role in assigning Local Priority and Color values. In ArubaOS-CX, you can configure trust at the global or Interface level. When the trust level is set at both levels, the interface level configuration takes precedence.

ArubaOS-CX supports three possible configurations: Trust CoS, Trust DSCP and Trust none. The first two options act as the name suggests - the switch accepts the incoming marking and uses it to determine class of service. In ArubaOS-CX 10.5 an interface can only support one trust setting at the time, meaning that you can set DSCP or CoS but not both at the same time, on the same interface.

It is not always recommended to trust ingress markings, especially if traffic is received from non-trusted devices. In these situations you can protect the network by ignoring the incoming marking setting the trust setting to none.



Global and Interface Trust

Untrusted Edge

DSCP:46 CoS:7

1/1/1

trust none

AOS-CX switch

DSCP:46 CoS:7

Prioritization

LP: 0
Color: Green

Derived from 0 in CoS map (user-configurable)

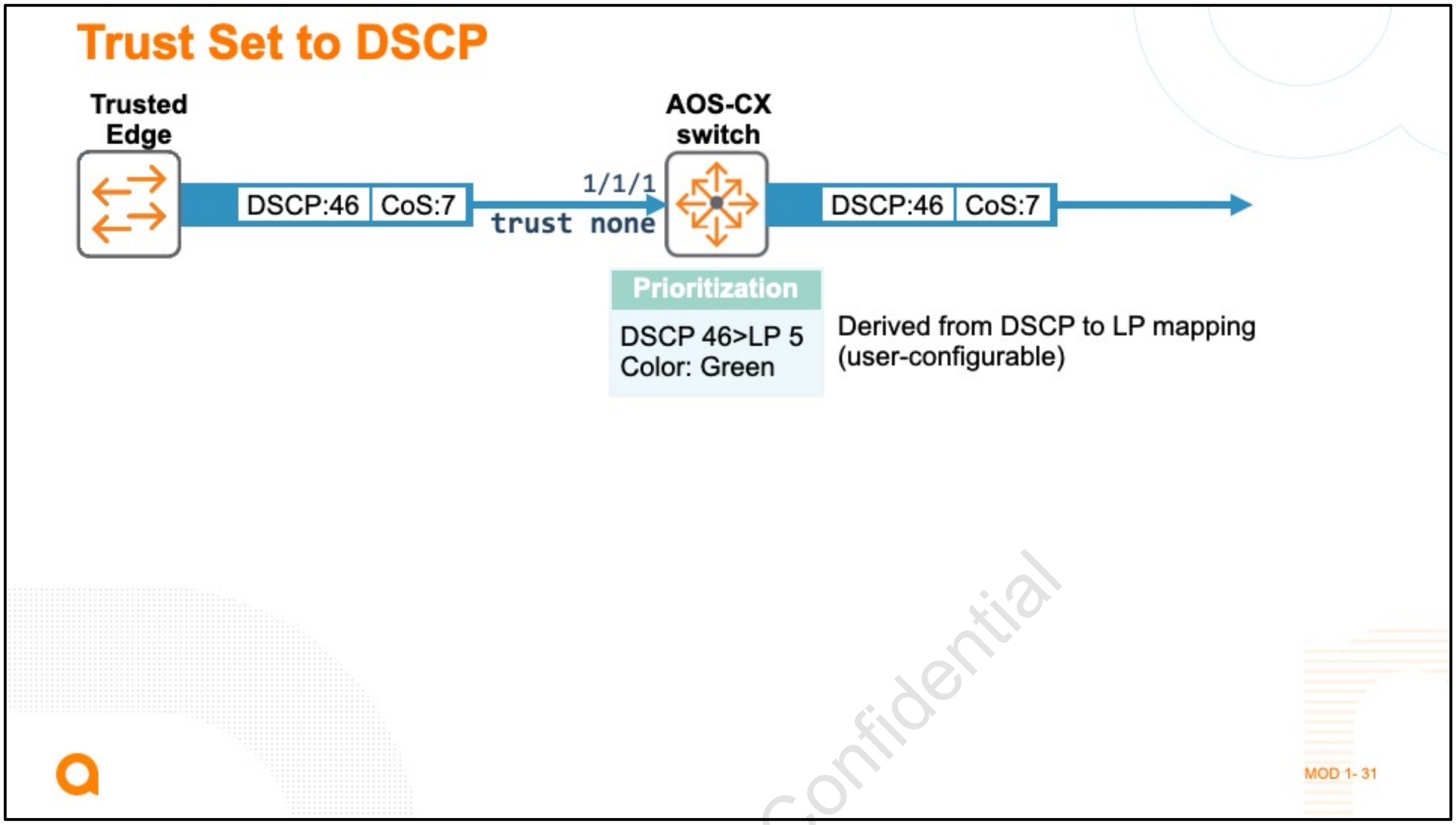
MOD 1- 30

If a port has no setting, the port takes its setting from the global setting. Otherwise, the port setting, including "none," overrides the global setting. In this way, you could globally not trust incoming marks, but trust DSCP on certain ports. You could globally trust DSCP, but use CoS instead on certain ports, and so on.

By default, the global setting is "none," and all ports have no trust setting. This means that all ports use the global setting of "none" - the switch ignores any markings in ingress frames and assigns all traffic the LP and color associated with CoS 0 in the CoS map. By default, this is LP=0 and color=green. The switch preserves any incoming DSCP when it forwards the traffic, as well as an incoming CoS value if the forwarded traffic is tagged. However, these values have no effect on how the switch queues the traffic.

RFC 2697 defines the use of colors with DSCP codes. A Single Rate Three Color Marker (srTCM) meters an IP packet stream and marks its packets either green, yellow, or red. Marking is based on a Committed Information Rate (CIR) and two associated burst sizes, a Committed Burst Size (CBS) and an Excess Burst Size (EBS). A packet is marked green if it doesn't exceed the CBS, yellow if it does exceed the CBS, but not the EBS, and red otherwise. The srTCM is useful, for example, for ingress policing of a service, where only the length, not the peak rate, of the burst determines service eligibility. See RFC 2697 for more information: https://www.ietf.org/rfc/rfc2697.txt. The use of colors is optional.

Note. It is important to clarify that a port not having a trust configuration is different from the port having a configuration of none. If a port has no setting, the port takes its setting from the global setting. Otherwise, the port setting, including "none," overrides the global setting.

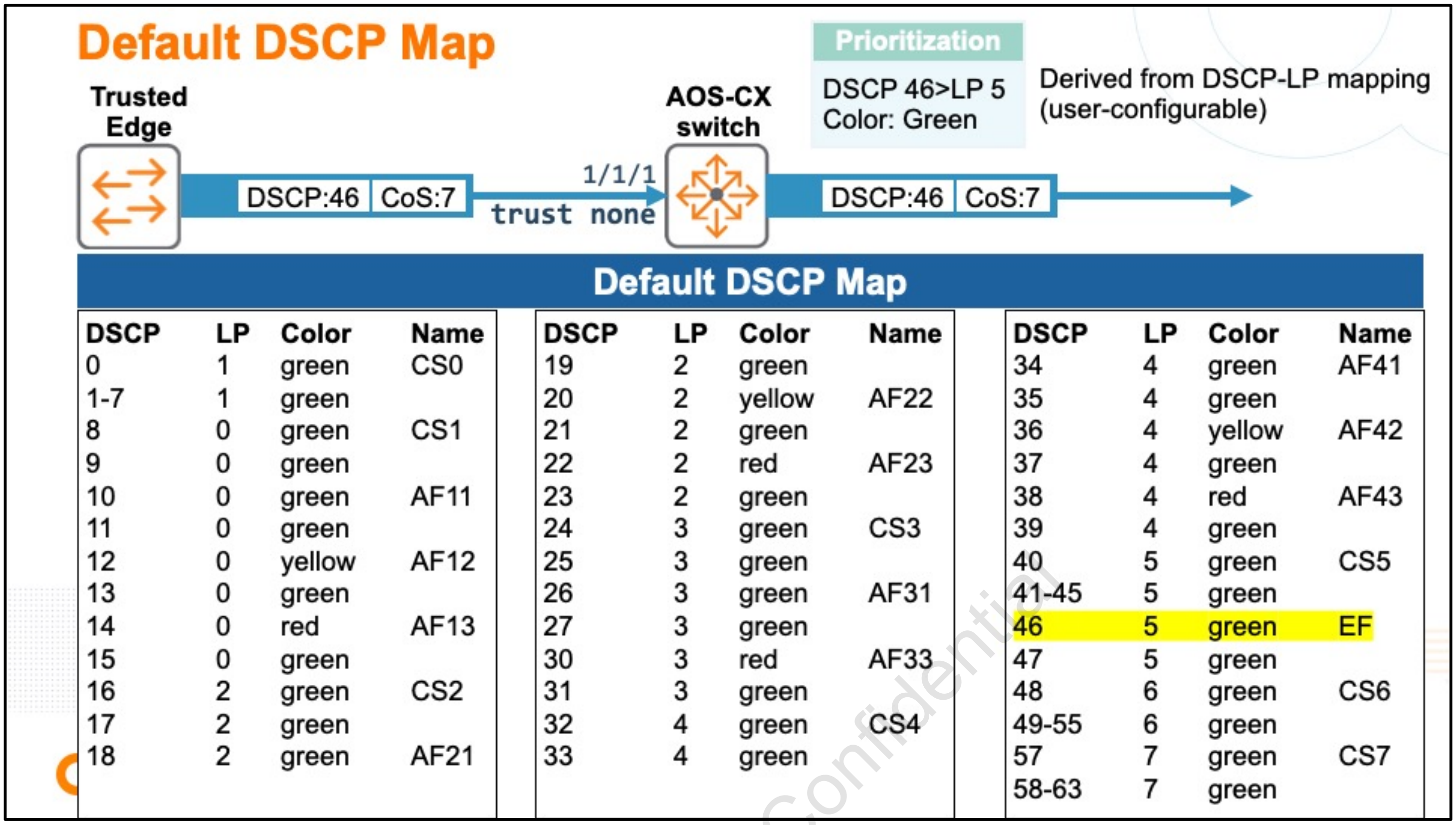


Trust Set to DSCP

DSCP 46>LP 5
Color: Green

Derived from DSCP to LP mapping (user-configurable)

Many networks classify traffic at the edge, so it can be useful for the AOS-CX switch to use the already applied classifications to determine the correct priority for the traffic. If other network infrastructure devices are using DSCP to classify traffic, set the global or port trust mode to dscp. When a port has a dscp trust setting, whether that is set on the port or taken from the global setting, the switch looks at the incoming DSCP in traffic's IP headers and uses its DSCP map to assign the traffic an LP and color.



Default DSCP Map

Prioritization

DSCP 46>LP 5
Color: Green

Derived from DSCP-LP mapping (user-configurable)

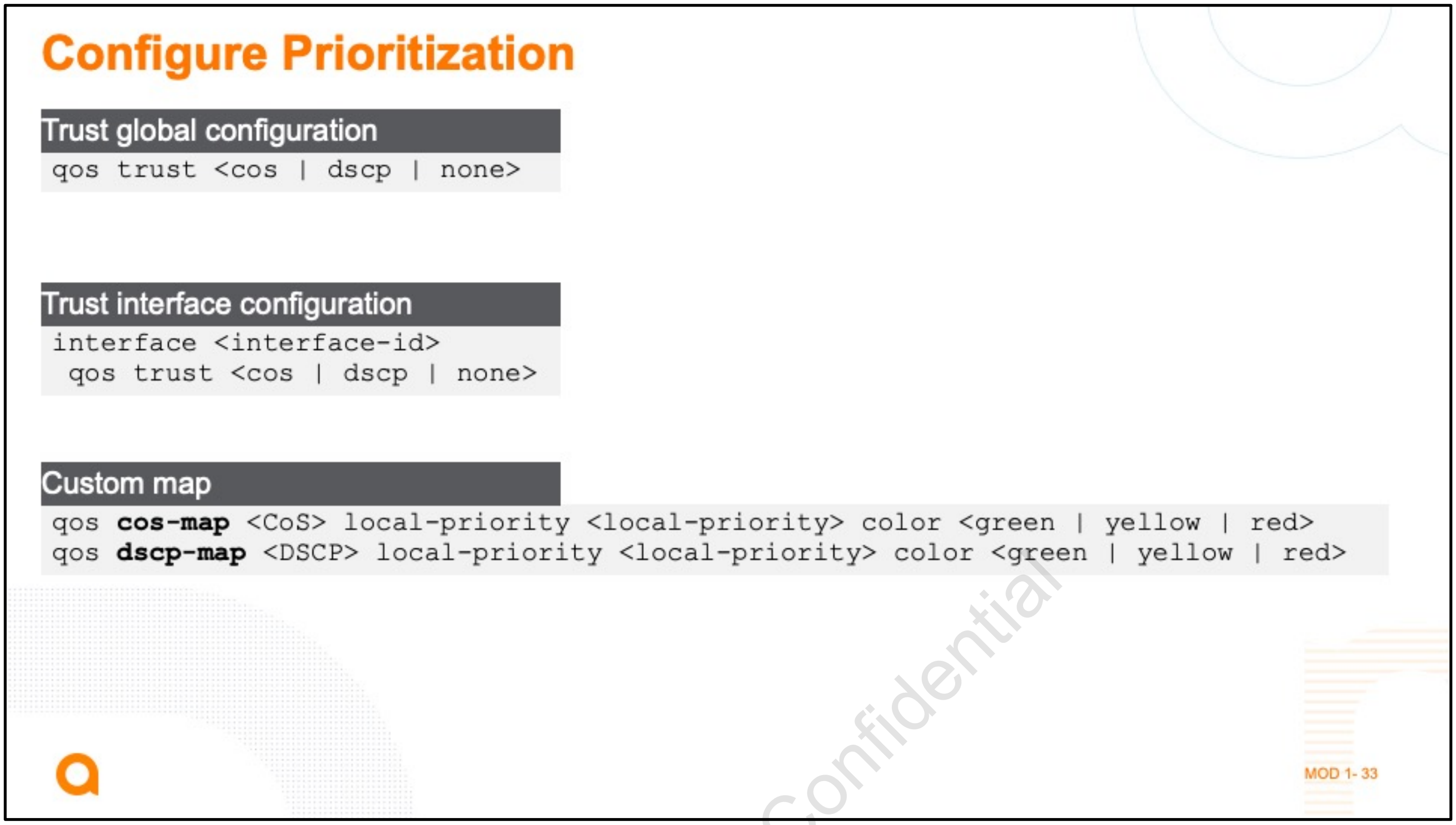
Default DSCP Map

DSCP	LP	Color	Name
0	1	green	CS0
1-7	1	green	
8	0	green	CS1
9	0	green	
10	0	green	AF11
11	0	green	
12	0	yellow	AF12
13	0	green	
14	0	red	AF13
15	0	green	
16	2	green	CS2
17	2	green	
18	2	green	AF21
19	2	green	
20	2	yellow	AF22
21	2	green	
22	2	red	AF23
23	2	green	
24	3	green	CS3
25	3	green	
26	3	green	AF31
27	3	green	
30	3	red	AF33
31	3	green	
32	4	green	CS4
33	4	green	
34	4	green	AF41
35	4	green	
36	4	yellow	AF42
37	4	green	
38	4	red	AF43
39	4	green	
40	5	green	CS5
41-45	5	green	
46	5	green	EF
47	5	green	
48	6	green	CS6
49-55	6	green	
57	7	green	CS7
58-63	7	green	

Here's the default DSCP Map, which follows the standards and works in many circumstances.

In this example, the map has an entry for DSCP 46, which assigns the incoming traffic LP 5 and color green. However, you can customize this map, setting your own local priority, color, and name for each DSCP. Your goal in customizing the map should be to match the system the devices use from where the switch is receiving the incoming traffic. For example, you may want to change DSCP 46, used for VoIP traffic, to LP 7 if the company wants to give this traffic highest priority.

Note that, if the incoming traffic has a CoS value as well as a DSCP, the switch preserves this value in the outgoing frame, if tagged. However, the CoS value and map have no effect on the traffic's LP and color. Those derive from the DCSP only.



Configure Prioritization

Trust global configuration

```
qos trust <cos | dscp | none>
```

Trust interface configuration

```
interface <interface-id>
 qos trust <cos | dscp | none>
```

Custom map

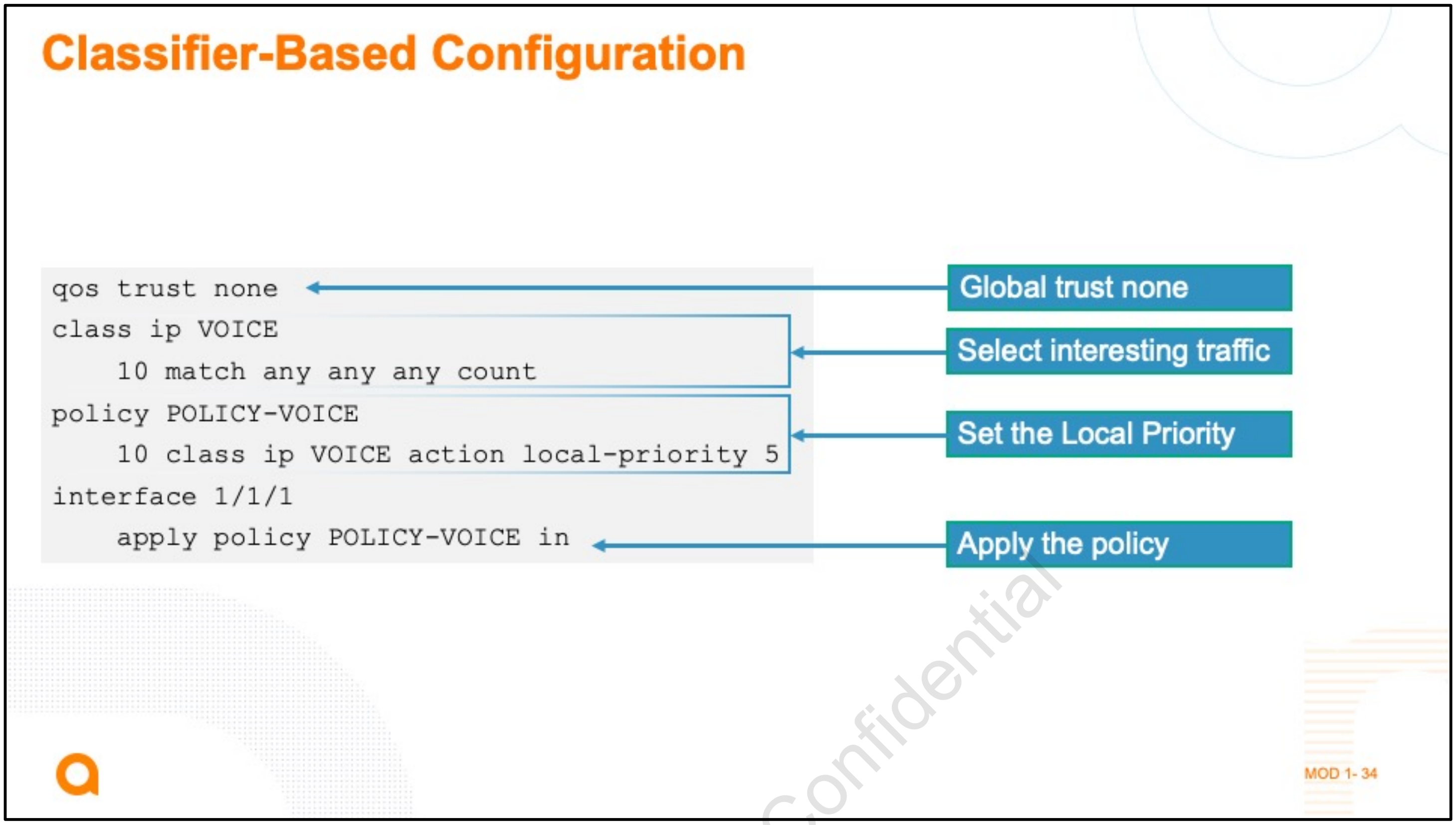
```
qos cos-map <CoS> local-priority <local-priority> color <green | yellow | red>
qos dscp-map <DSCP> local-priority <local-priority> color <green | yellow | red>
```

MOD 1- 33

You can configure the prioritization stage as shown in the figure.

For the custom map has two options. Use the cos-map option to define the local priority assigned to incoming packets for a specific 802.1 VLAN priority code point (CoS) value. The CoS map values are used to mark incoming packets when QoS trust mode is set to cos.

Use the dscp-map option to defines the local priority assigned to incoming packets for a specific IP differentiated services code point (DSCP) value. The DSCP map values are used to prioritize incoming packets when QoS trust mode is set to dscp.

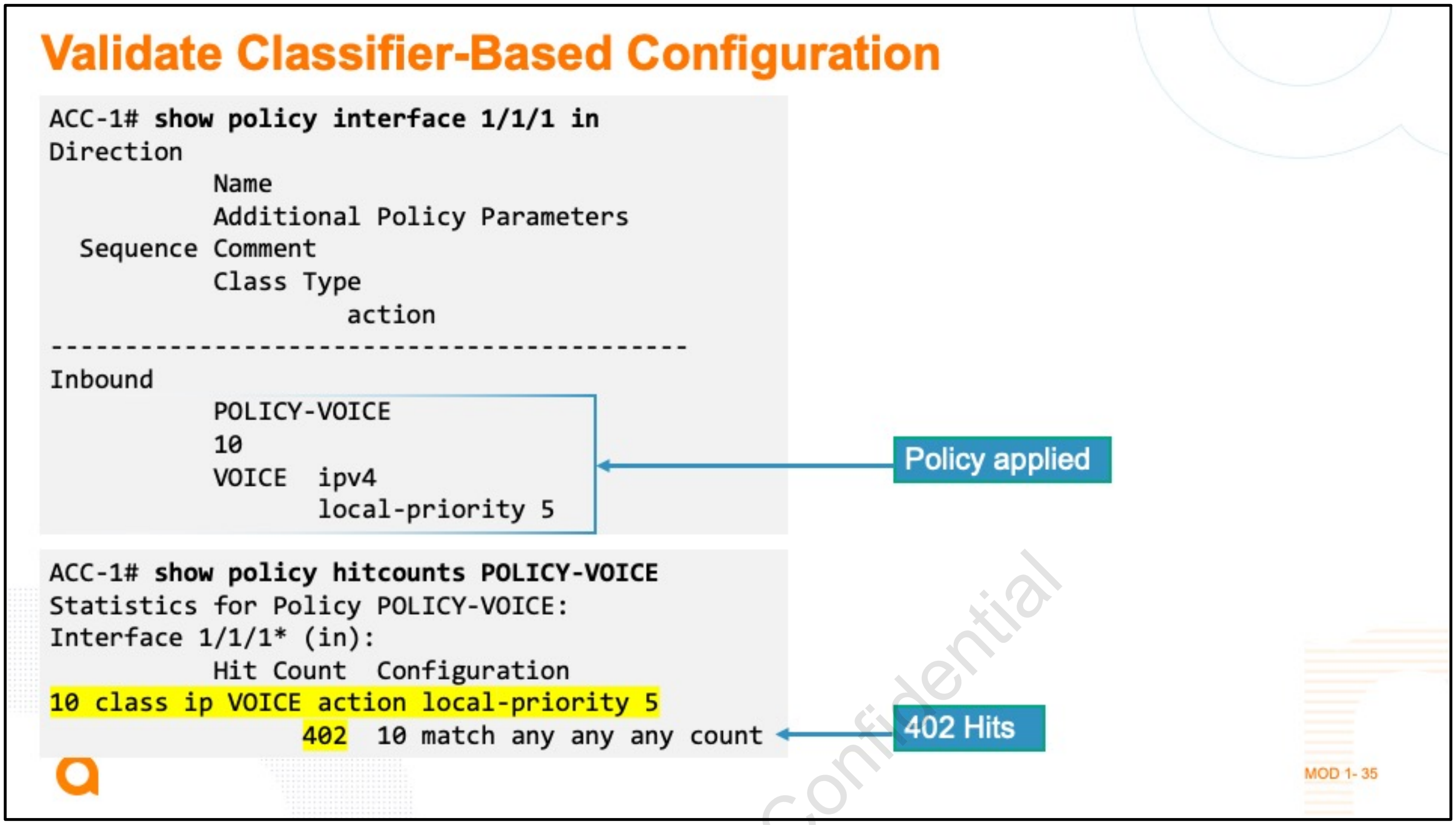


Classifier-Based Configuration

```
qos trust none
class ip VOICE
    10 match any any any count
policy POLICY-VOICE
    10 class ip VOICE action local-priority 5
interface 1/1/1
    apply policy POLICY-VOICE in
```

Global trust none
Select interesting traffic
Set the Local Priority
Apply the policy
MOD 1- 34

This example shows how you can use a classifier-based policy configuration to assign the Local Priority.



Validate Classifier-Based Configuration

```
ACC-1# show policy interface 1/1/1 in
Direction
           Name
           Additional Policy Parameters
  Sequence Comment
           Class Type
                    action
-------------------------------------------
Inbound
           POLICY-VOICE
           10
           VOICE  ipv4
                  local-priority 5
```

Policy applied

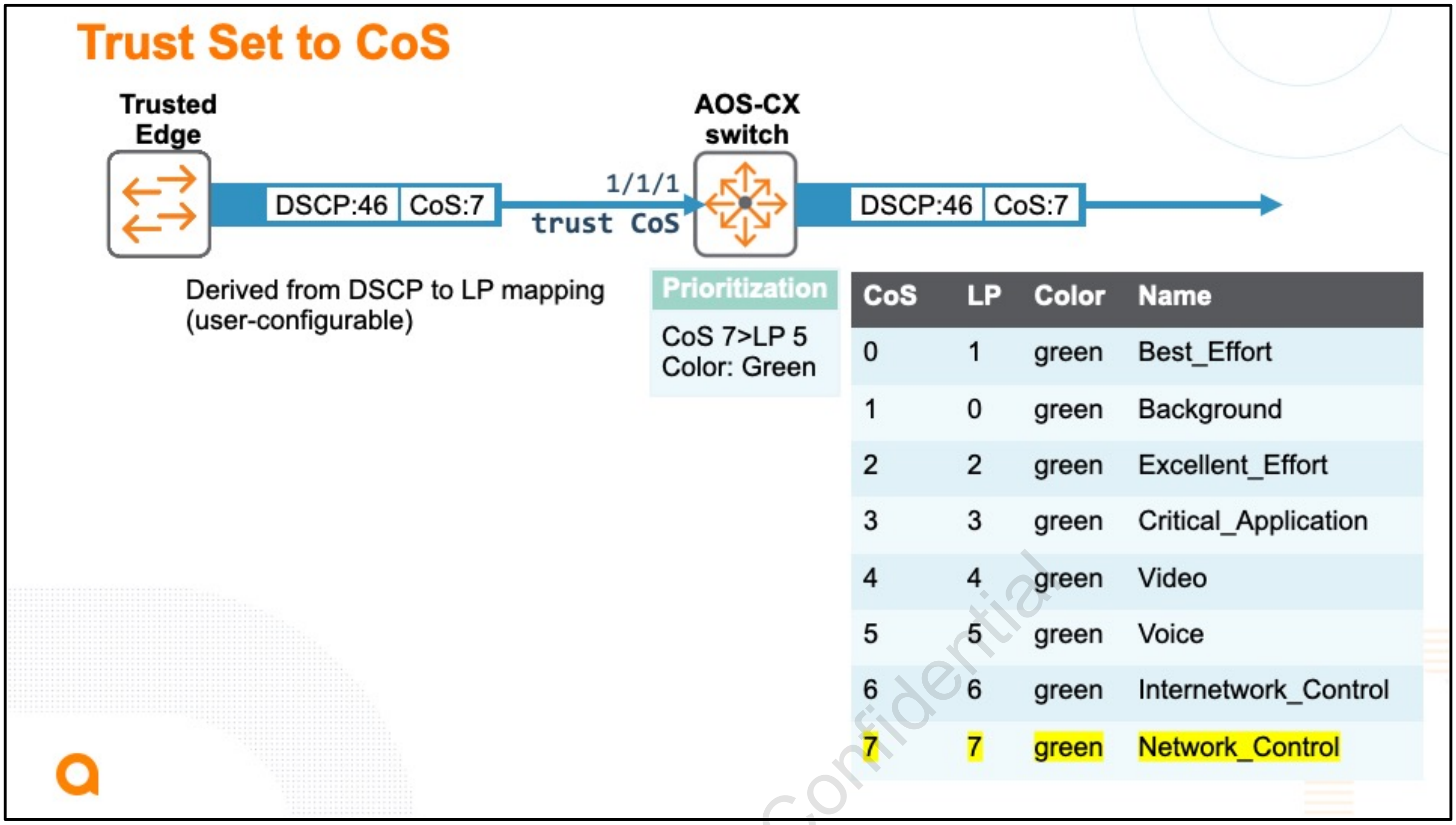
```
ACC-1# show policy hitcounts POLICY-VOICE
Statistics for Policy POLICY-VOICE:
Interface 1/1/1* (in):
           Hit Count  Configuration
10 class ip VOICE action local-priority 5
                 402  10 match any any any count
```

402 Hits

MOD 1- 35

This example shows how to validate the classifier-based policy configuration from the previous example.

Use the show policy interface <interface-id> in command to verify if the policy was correctly set on the interface. The command show policy hitcounts <Policy-name> displays the number of hits counts that the policy received.



Trust Set to CoS

Trusted Edge

DSCP:46 CoS:7

1/1/1

trust CoS

AOS-CX switch

DSCP:46 CoS:7

Derived from DSCP to LP mapping (user-configurable)

Prioritization

CoS 7>LP 5
Color: Green

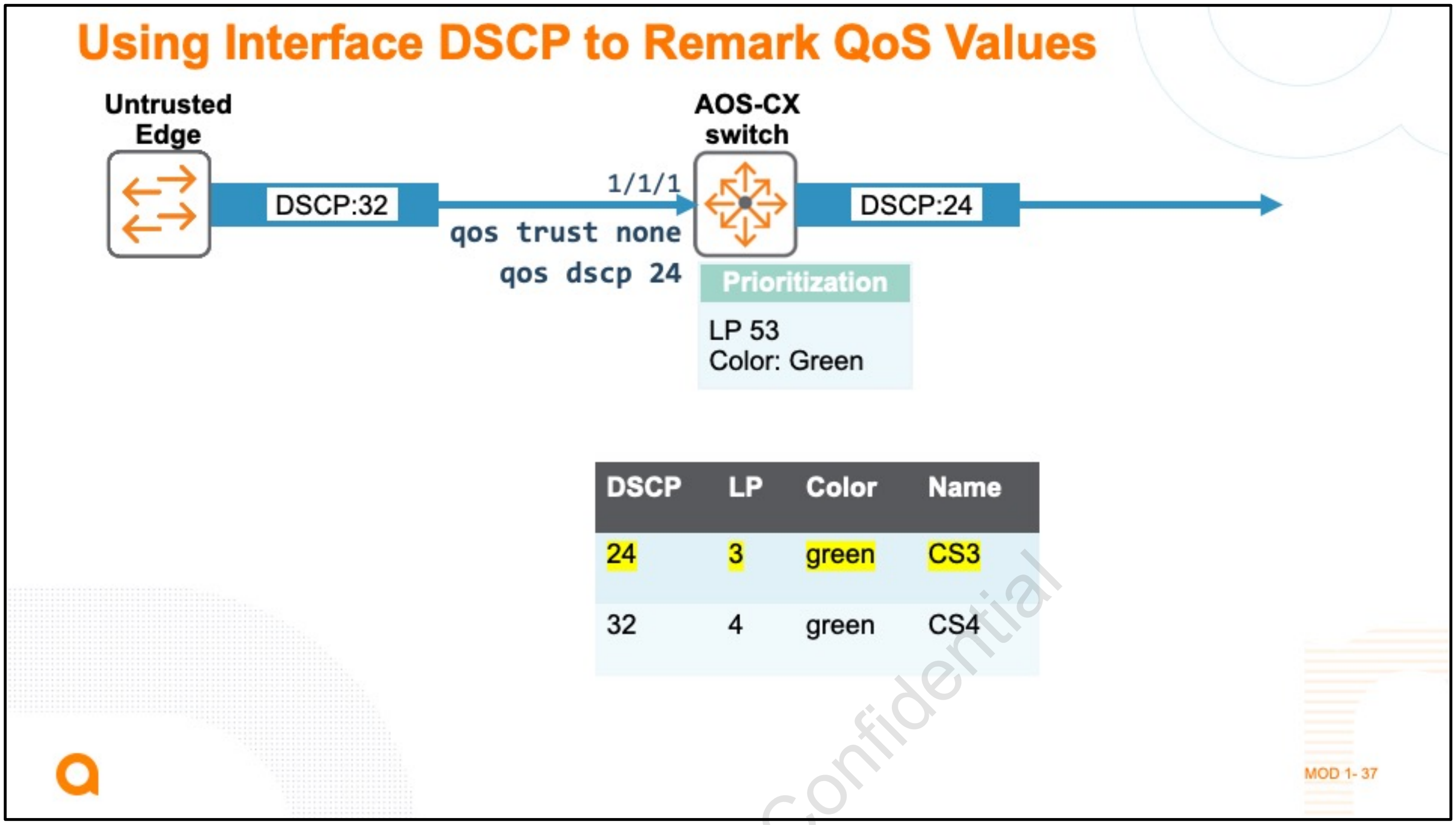
CoS	LP	Color	Name
0	1	green	Best_Effort
1	0	green	Background
2	2	green	Excellent_Effort
3	3	green	Critical_Application
4	4	green	Video
5	5	green	Voice
6	6	green	Internetwork_Control
7	7	green	Network_Control

Instead of using DSCP to assign the LP and color, you can use the CoS value by configuring ports to trust to cos. You can set the trust globally and leave ports without individual trust settings, or you can set the trust setting on individual ports. Keep in mind that the CoS is carried in the VLAN tag, so you should not use this setting for ports that receive native VLAN traffic or operate in route-only mode. All traffic that arrives without a VLAN tag is considered as having CoS value 0. In general, you might use the cos trust setting if port(s) receive non-IP traffic or if the rest of the network uses CoS in preference to DSCP.

This is an example of the default CoS map with the LP and color values associated with each CoS value.

This table provides the traffic LP 7 and color green. AOS-CX switches use the 802.1Q-2014 recommendations in which CoS 1 has the lowest priority, LP 0 and CoS 0 has the next lowest, LP 1. The rest of the CoS values are mapped to an LP with the same value. Again, you can adjust the map to assign each CoS value the LP and color that you desire.

When a port trusts CoS, whether based on the global or port setting, incoming DSCP has no effect on the traffic's LP and color. However, the switch does preserve this DSCP in forwarding the traffic.



Using Interface DSCP to Remark QoS Values

```
qos trust none
qos dscp 24
```

Prioritization

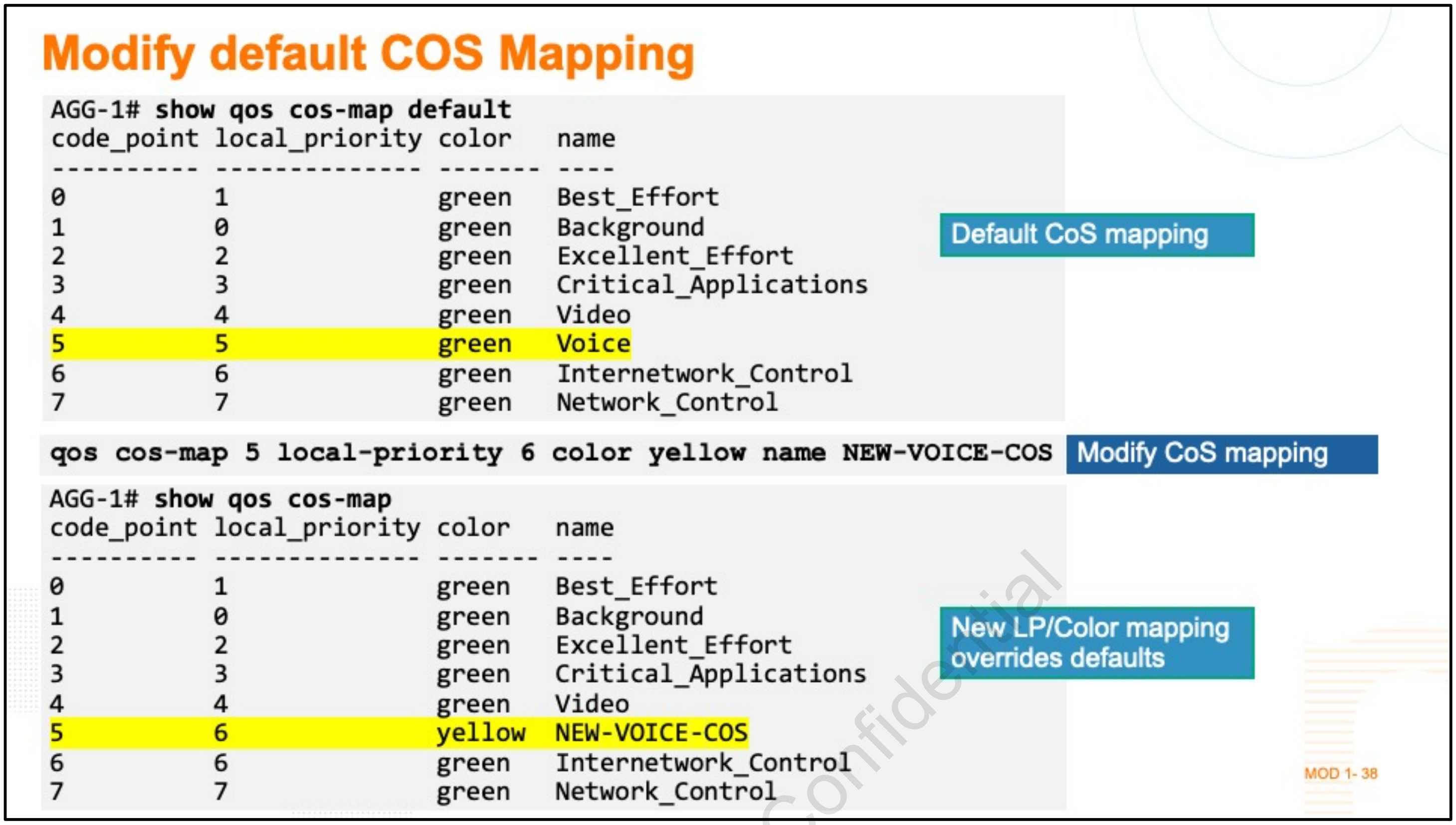
LP 53
Color: Green

DSCP	LP	Color	Name
24	3	green	CS3
32	4	green	CS4

MOD 1- 37

Rather than use the incoming priority to assign the LP, AOS-CX switches can also classify traffic on their own. You should use this approach when connected devices are not capable of or not trusted to assign the correct values. The simplest approach is to assign a DSCP value to the Ethernet interface or LAG. When you use this approach, the trust setting on the interface must be none. Then the port overwrites any incoming DSCP with the specified DSCP value on all traffic that arrives on the port. The switch uses that new value and the DSCP map to assign the traffic its LP and color.

This does not affect the CoS, and also does not affect the prioritization. In this example, traffic is untagged (native VLAN) and has no CoS.



Modify default COS Mapping

```
AGG-1# show qos cos-map default
code_point local_priority color   name
---------- -------------- ------- ----
0          1              green   Best_Effort
1          0              green   Background
2          2              green   Excellent_Effort
3          3              green   Critical_Applications
4          4              green   Video
5          5              green   Voice
6          6              green   Internetwork_Control
7          7              green   Network_Control
```

Default CoS mapping

```
qos cos-map 5 local-priority 6 color yellow name NEW-VOICE-COS
```

Modify CoS mapping

```
AGG-1# show qos cos-map
code_point local_priority color   name
---------- -------------- ------- ----
0          1              green   Best_Effort
1          0              green   Background
2          2              green   Excellent_Effort
3          3              green   Critical_Applications
4          4              green   Video
5          6              yellow  NEW-VOICE-COS
6          6              green   Internetwork_Control
7          7              green   Network_Control
```

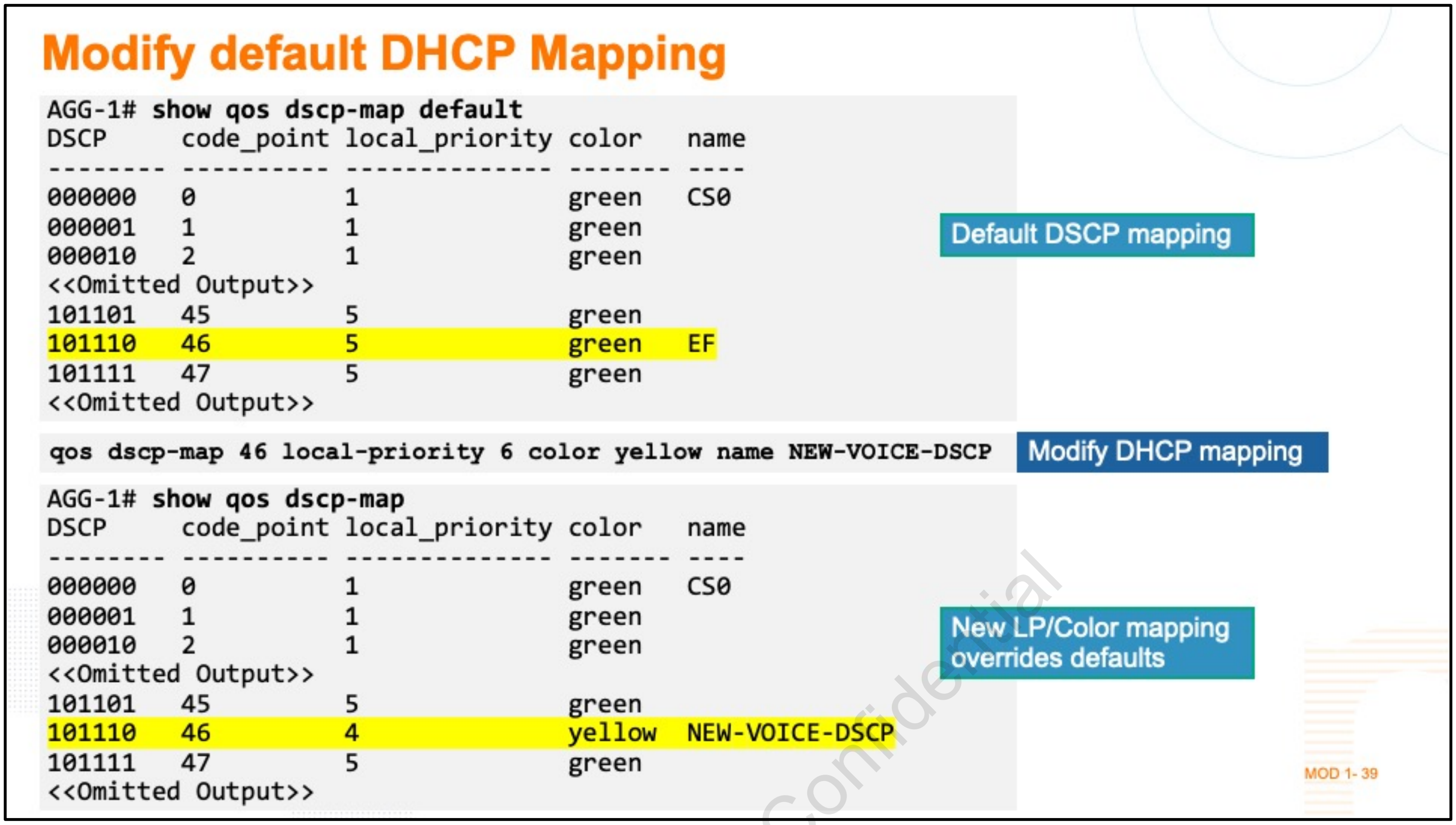
New LP/Color mapping overrides defaults

MOD 1- 38

Use the show qos cos-map default command to verify the default mapping for incoming CoS marks.

You can customize the ArubaOS-CX default mappings. The command changes one row in the map, specifying the desired LP, color, and name to associate with a specific CoS value. Specifying the LP is mandatory, but the color and name are optional. If you do not specify a color, the color is set to green. If you do not specify a name, the row in the map has no name.

In this example, suppose you want to modify the default mapping for the voice marks using the command shown in the figure. This action overrides the default mapping, which you can verify with the show qos cos-map.



Modify default DHCP Mapping

```
AGG-1# show qos dscp-map default
DSCP     code_point local_priority color   name
-------- ---------- -------------- ------- ----
000000   0          1              green   CS0
000001   1          1              green
000010   2          1              green
<<Omitted Output>>
101101   45         5              green
101110   46         5              green   EF
101111   47         5              green
<<Omitted Output>>
```

Default DSCP mapping

```
qos dscp-map 46 local-priority 6 color yellow name NEW-VOICE-DSCP
```

Modify DHCP mapping

```
AGG-1# show qos dscp-map
DSCP     code_point local_priority color   name
-------- ---------- -------------- ------- ----
000000   0          1              green   CS0
000001   1          1              green
000010   2          1              green
<<Omitted Output>>
101101   45         5              green
101110   46         4              yellow  NEW-VOICE-DSCP
101111   47         5              green
<<Omitted Output>>
```

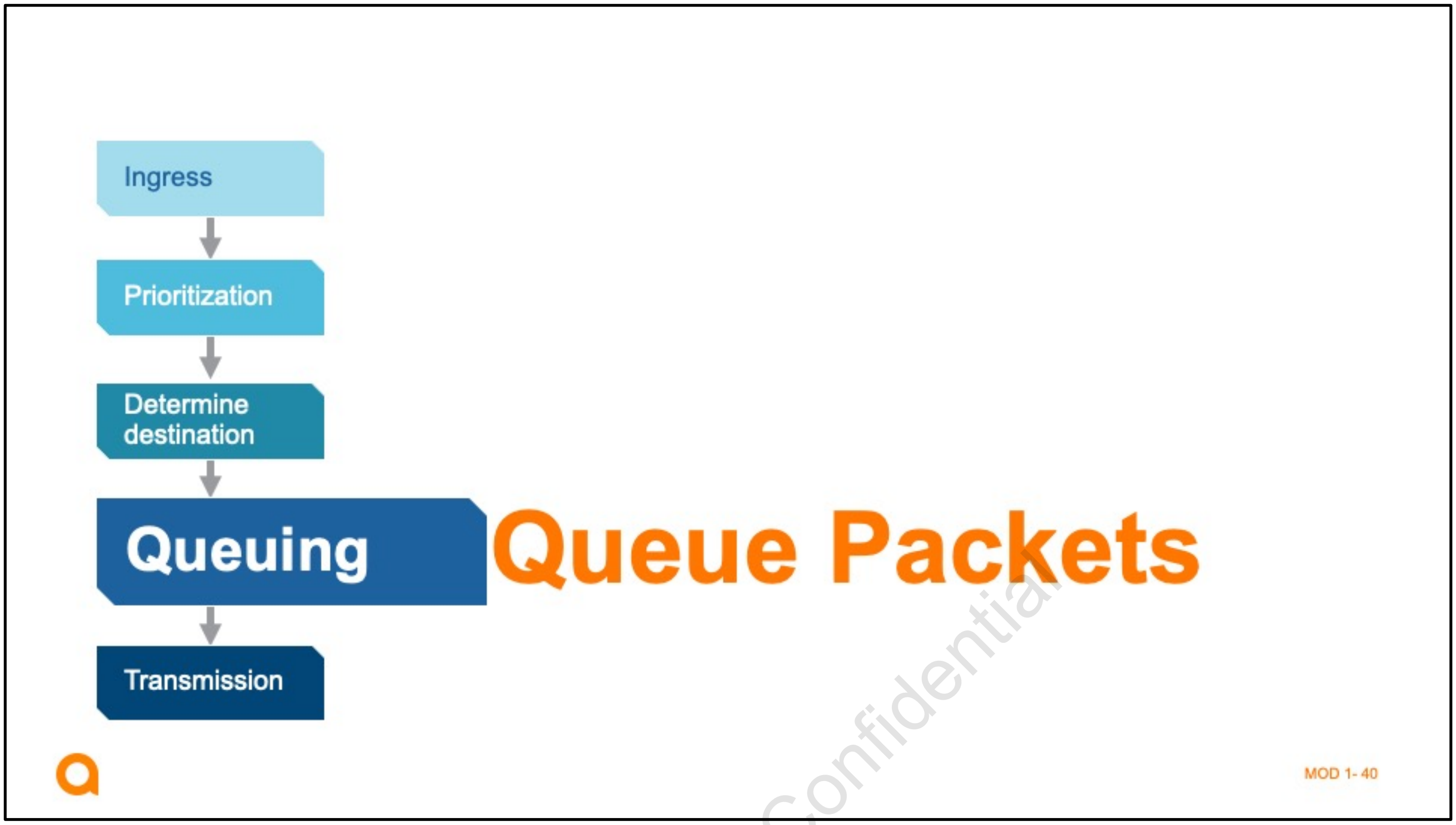
New LP/Color mapping overrides defaults

MOD 1- 39

Use the show qos dscp-map default command to verify the default mapping for incoming DSCP marks.

Customize these defaults with the command shown in the figure. The command changes one row in the map, specifying the desired LP, color, and name to associate with a specific DSCP value. Specifying the LP is mandatory, but the color and name are optional. If you do not specify a color, the color is set to green. If you do not specify a name, the row in the map has no name.

This action overrides the default mapping, which you can verify with the show qos dscp-map command.



Ingress

Prioritization

Determine destination

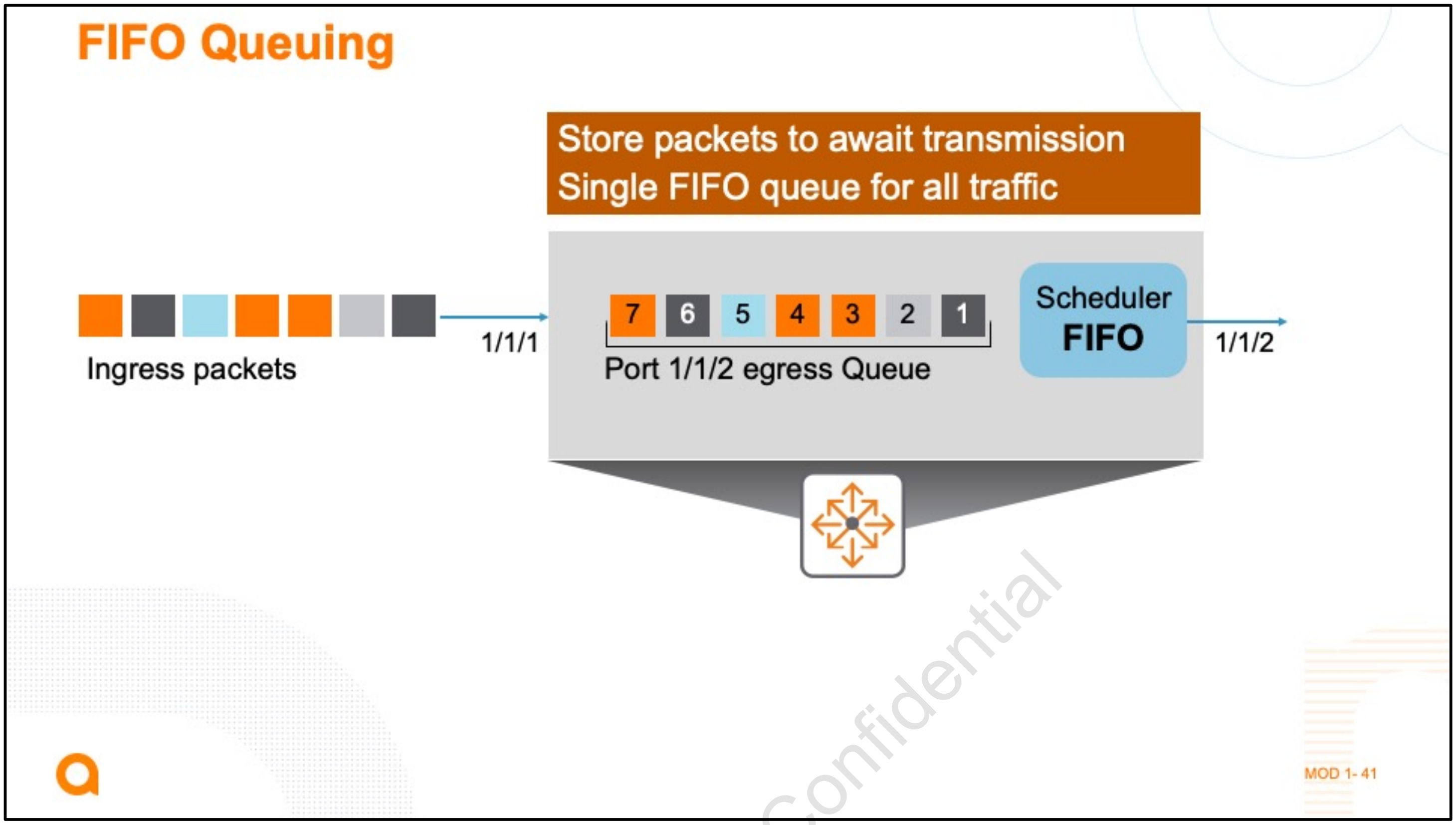
Queuing

Queue Packets

Transmission

MOD 1- 40

Now we move to packet queueing.



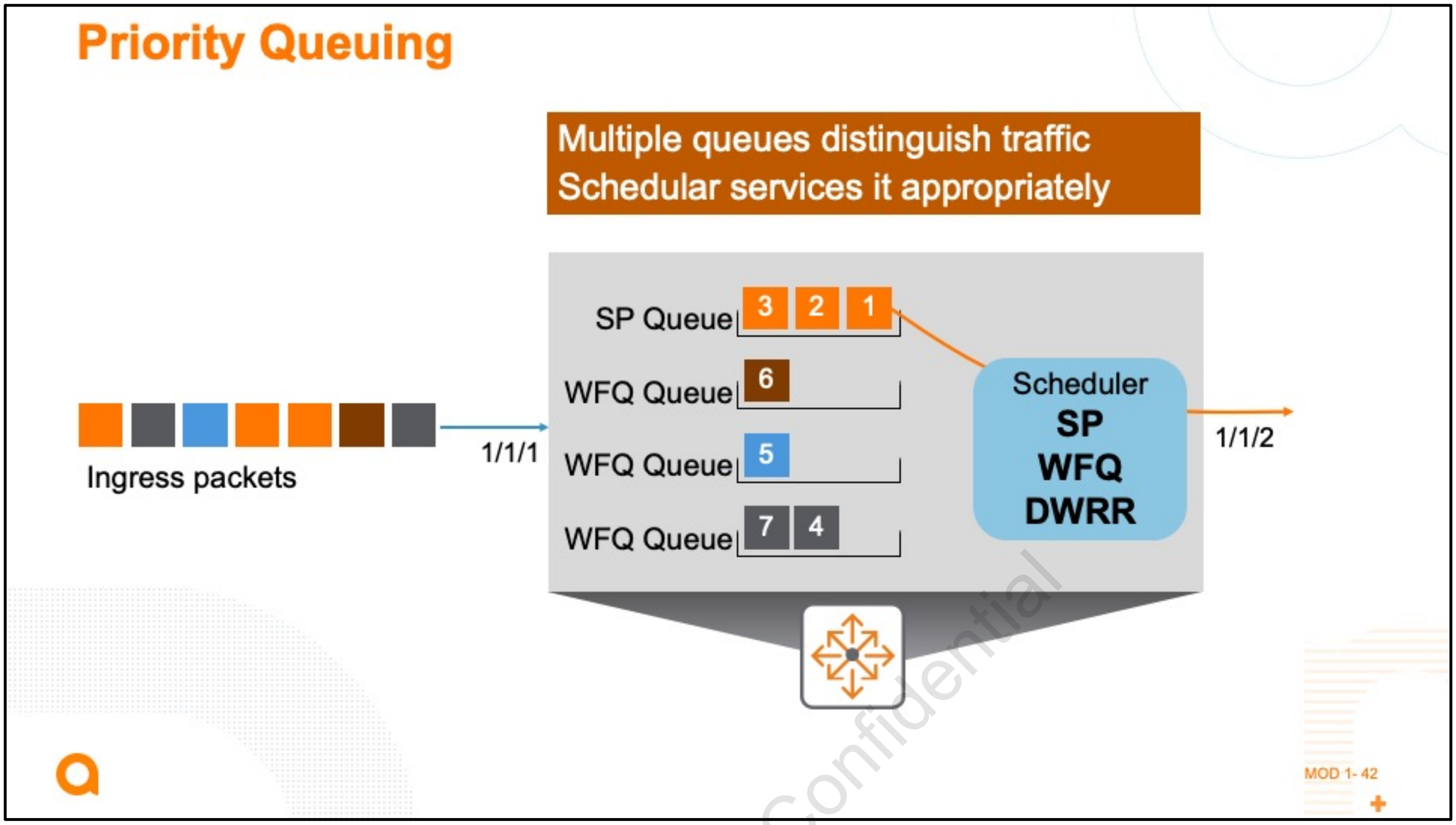
FIFO Queuing

Store packets to await transmission
Single FIFO queue for all traffic

As a switch prepares to transmit traffic, it places the traffic in an egress queue for the egress port. Then the device must schedule packets for service – transmit them out the egress port. If the queue works in a First In First Out (FIFO) fashion, packets are sent in the order received, as shown in the figure. Using a single queue offers no advantage to sensitive traffic.

Perhaps the orange VoIP packets are stuck waiting behind large FTP packets. The FTP traffic would not be affected by an extra 100ms of delay, but the voice quality suffers because of this delay.

The main parameter that defines the queue is its length, which is how many packets it can store. In ArubaOS-CX each interface has an egress buffer up to 1MB.



Priority Queuing

Multiple queues distinguish traffic
Schedular services it appropriately

To prioritize traffic, the switch can use a queuing mechanism that sorts various packets into multiple queues. AOS-CX switches support up to eight egress queues per-port.

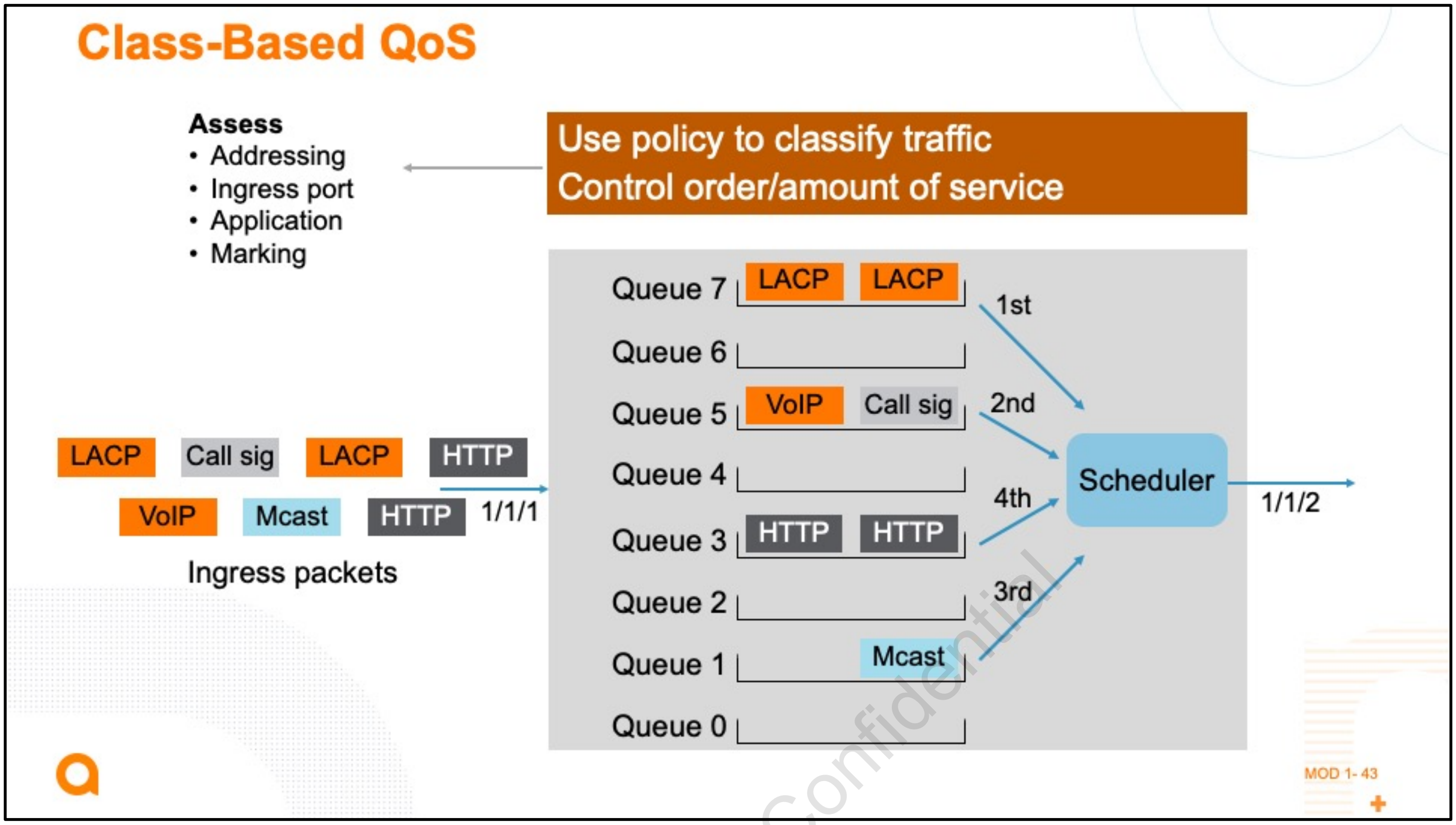
The figure shows a simplified scenario with four queues.

| The Orange VoIP packets are placed in one queue.

| Various other traffic placed in other queues.

| Now that traffic is organized into queues, the port’s scheduling mechanism selects the next packet for forwarding. Scheduling works in tandem with queuing, ensuring that packets in high-priority queues are transmitted more promptly than packets in other queues.

| AOS-CX switches support strict priority (SP) queuing and Weighted Fair Queuing (WFQ) or Deficit Weighted Round Robin (DWRR) queuing. Basically, the Strict Priority Queue is serviced until it is empty. Then it services the other queues using some variation of a round-robin, ensuring that all traffic types get their fair share of service.



Class-Based QoS

Assess
- Addressing
- Ingress port
- Application
- Marking

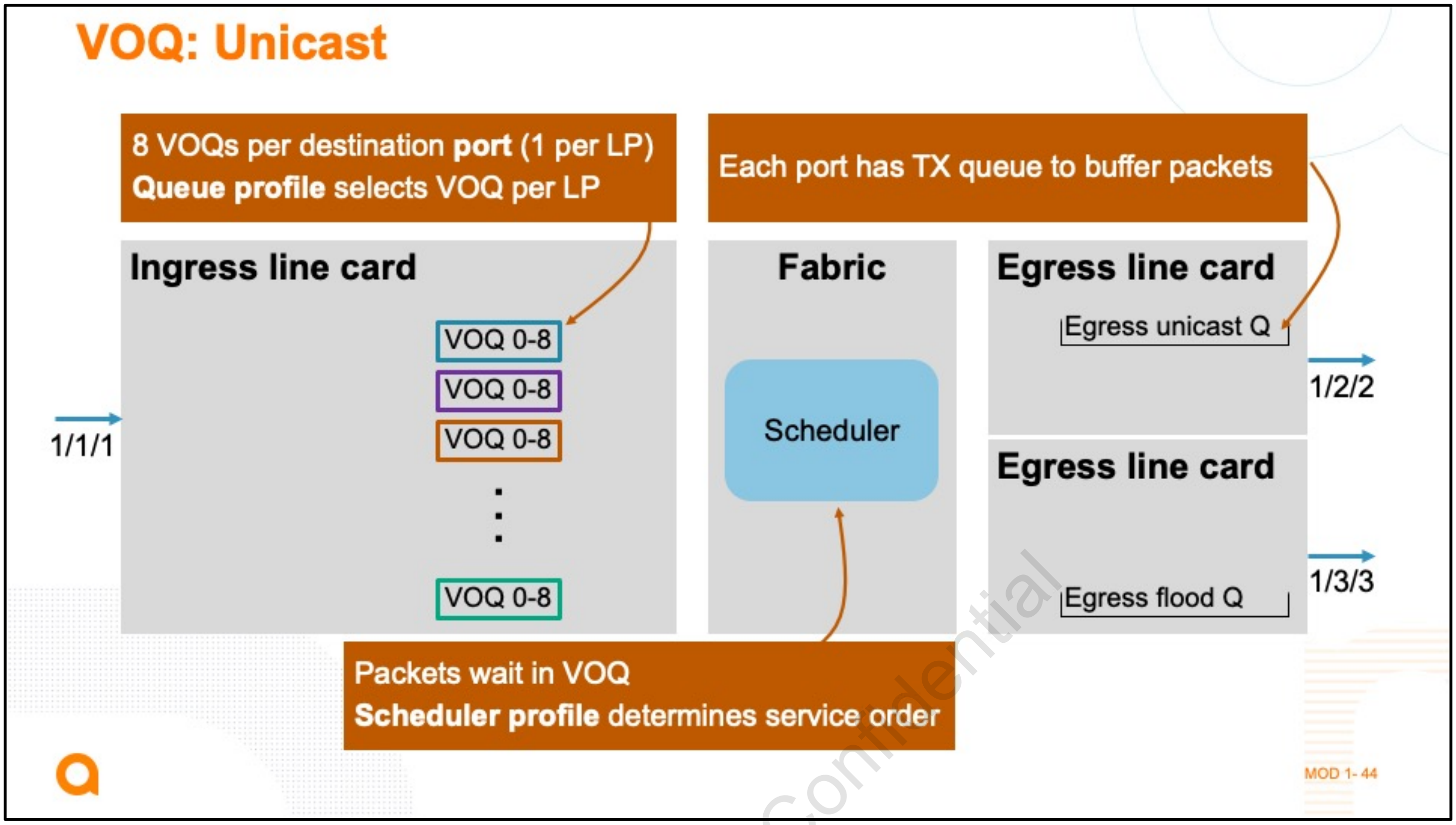
Use policy to classify traffic
Control order/amount of service

With QoS, every port has eight egress queues, numbered 0 to 7. You can define class-based policies and apply them to appropriate interfaces.

| Based on this policy, the switch classifies each ingress packet. This classification could be based on ingress interface, IP addresses, MAC addresses, pre-existing ingress marking, and more. That priority then determines the correct queue for the traffic.

The queue number is not relevant to how queues are serviced. Scheduling determines how the switch selects which queue forwards traffic at any given moment in a port's service cycle.

| Perhaps the scheduler services Queue 7 1st, then queue 5, then the multicast traffic in queue 1, and then the HTTP traffic in queue 3.



VOQ: Unicast

8 VOQs per destination **port** (1 per LP)
Queue profile selects VOQ per LP

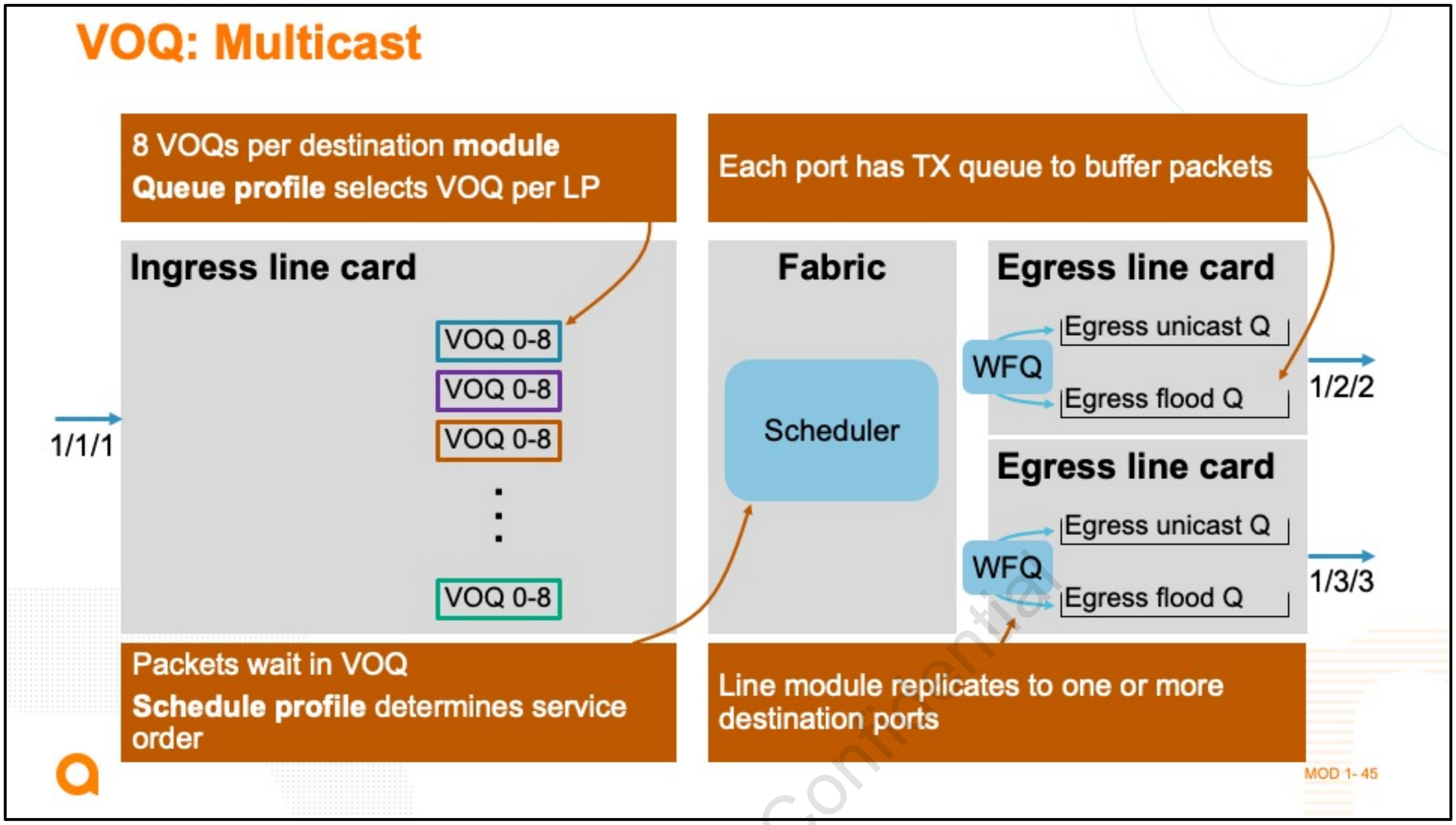
Each port has TX queue to buffer packets

Packets wait in VOQ
Scheduler profile determines service order

AOS-CX switches use queue and schedule profiles for making these decisions. The queue profile is global, while you can apply schedule profiles globally or to individual ports.

The switch uses a virtual output queue (VOQ) architecture where most packet buffering occurs on the ingress line module. Unicast traffic destined for one port uses different buffering and scheduling than BUM traffic destined to be flooded out multiple ports. The relative priority and the amount of packet transmission for these two types of traffic differ.

For unicast traffic, each line module contains eight VOQs for every destination port in the chassis (one per local priority). The queue profile determines which VOQ is used to buffer a local priority. The schedule profile determines the order of VOQ servicing. Unicast packets wait in VOQs until the scheduler selects them to cross the fabric. Each destination port has a shallow egress transmit queue that buffers unicast packets.



VOQ: Multicast

8 VOQs per destination **module**
Queue profile selects VOQ per LP

Each port has TX queue to buffer packets

Ingress line card

Fabric

Egress line card

VOQ 0-8
VOQ 0-8
VOQ 0-8
⋮
VOQ 0-8

Scheduler

WFQ
Egress unicast Q
Egress flood Q

Egress line card

WFQ
Egress unicast Q
Egress flood Q

1/1/1

1/2/2

1/3/3

Packets wait in VOQ
Schedule profile determines service order

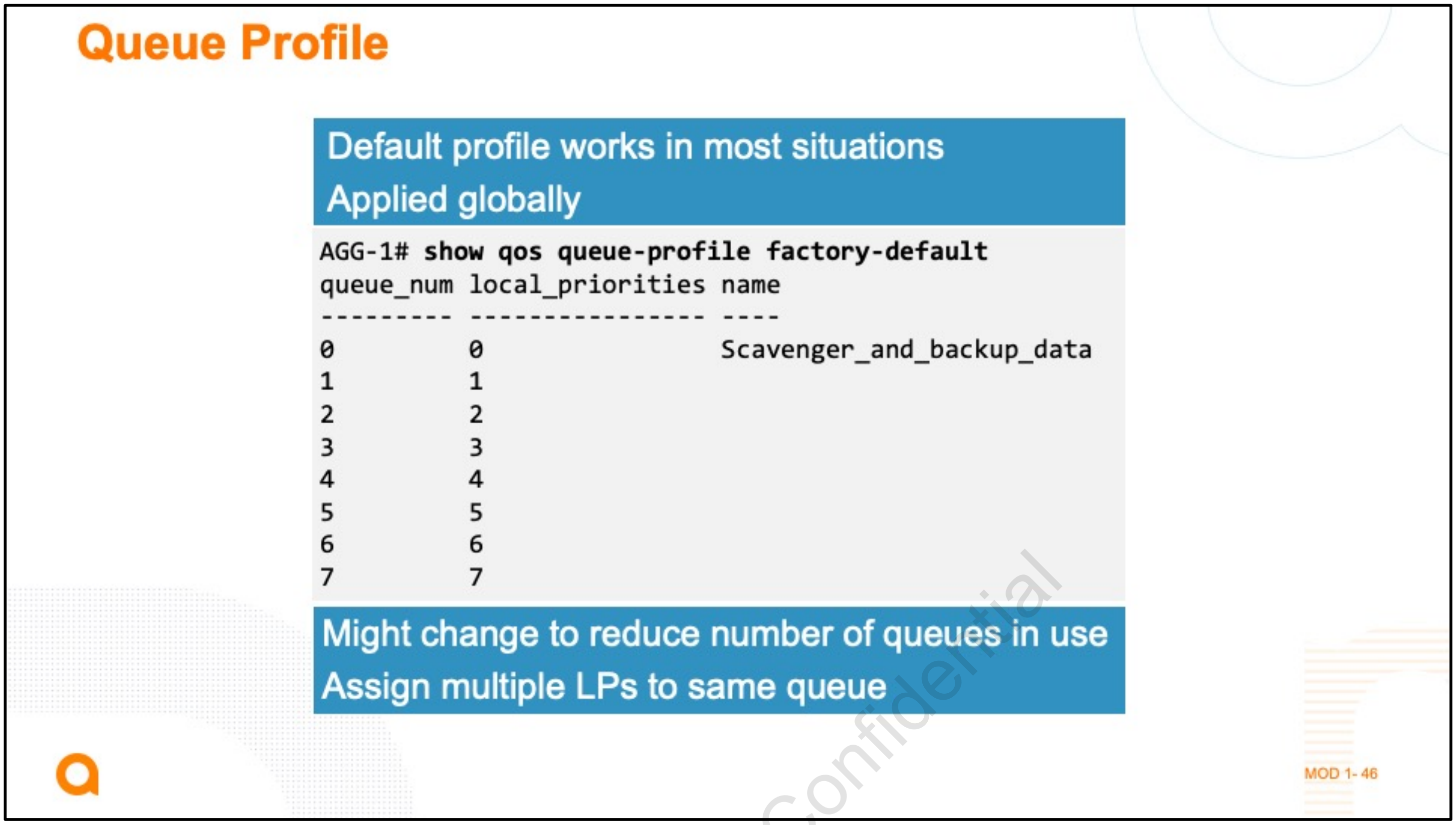
Line module replicates to one or more destination ports

MOD 1- 45

Broadcast, multicast, and unknown-unicast (BUM) packets, collectively called flooded traffic, use a separate path to the destination ports. Each line module contains eight VOQs per destination line module (including itself) to buffer traffic to be flooded (one VOQ per local priority). A copy of the packet to be flooded is buffered in VOQs for each destination line module. The queue profile determines which VOQ is used for each local priority.

Flooded traffic VOQs use a fixed strict schedule profile to determine the order of VOQ servicing. Flooded packets wait in VOQs until the scheduler selects them to cross the fabric. On the destination line module, they are replicated to one or more destination ports. Each destination port has a second shallow egress queue for replicated packets buffered for transmit.

A WFQ scheduler is used to select packets for transmission from the unicast and replicated traffic egress queues. When selecting packets between two non-empty queues, WFQ uses a fixed data weight of four for unicast traffic, and a weight of one for replicated traffic. As long as both queues are non-empty, replicated (flooded) packets comprise approximately 20% of the transmitted traffic, independent of the unicast scheduled percentages.



Queue Profile

Default profile works in most situations

Applied globally

```
AGG-1# show qos queue-profile factory-default
queue_num local_priorities name
--------- ---------------- ----
0         0                Scavenger_and_backup_data
1         1
2         2
3         3
4         4
5         5
6         6
7         7
```

Might change to reduce number of queues in use

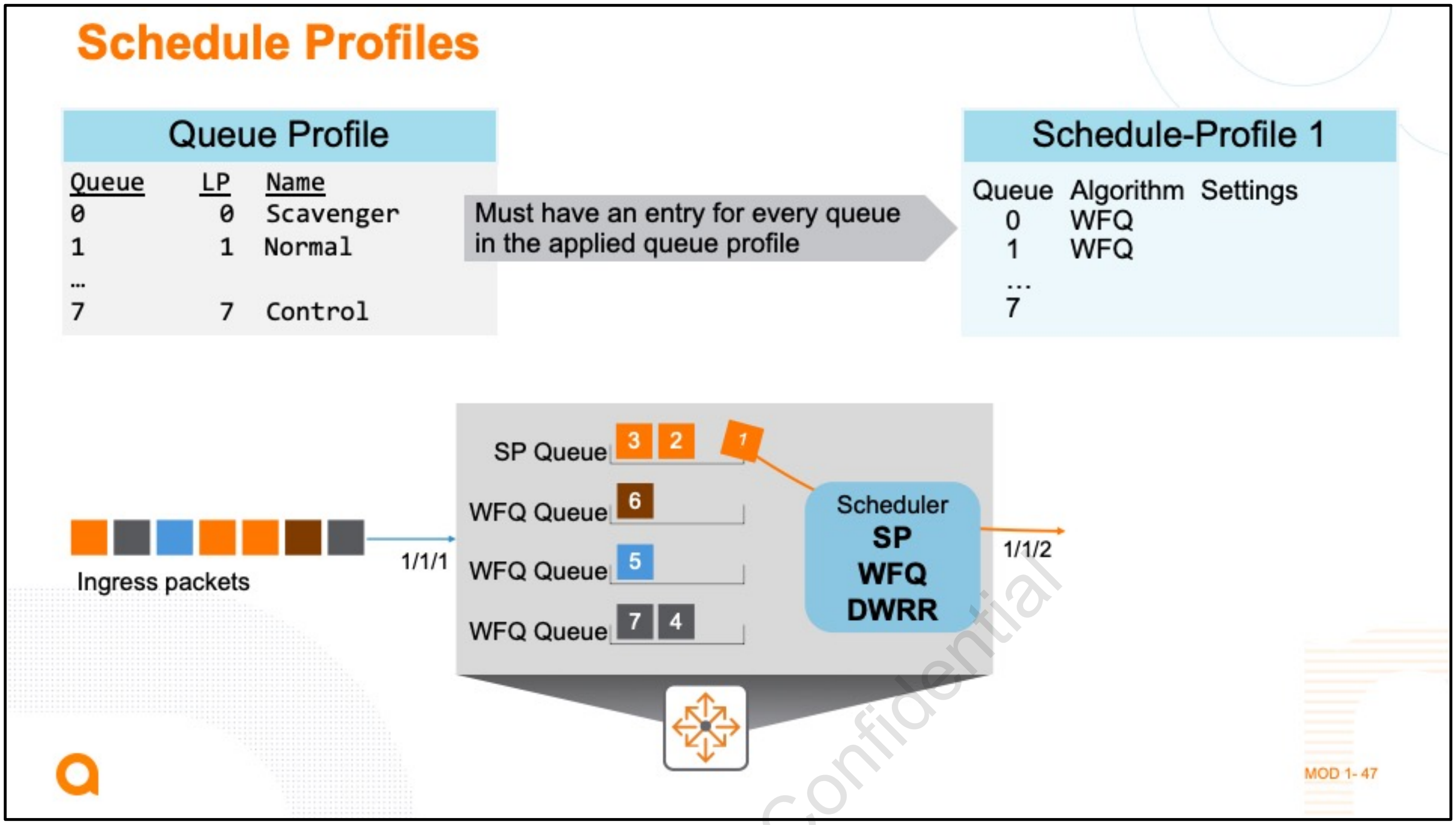
Assign multiple LPs to same queue

MOD 1- 46

The queue profile translates the LP to the priority queue from the traffic is forwarded. By default, each LP is mapped to a queue with the same numerical value. As shown in the figure Queue number 0 maps to LP 0, and so on.

You will typically keep this profile since you can adjust how QoS marks map to LP in the CoS and DSCP maps. However, you can also create a new profile with customized mappings.

For example, you may want to reduce the number of queues in use, which can simplify when your network is not classifying traffic granularly and some LPs are unused. You can assign more than one LP to the same queue to ensure that even LPs that you do not think you will use map to a queue.



Schedule Profiles

Queue Profile

Queue	LP	Name
0	0	Scavenger
1	1	Normal
…		
7	7	Control

Must have an entry for every queue in the applied queue profile

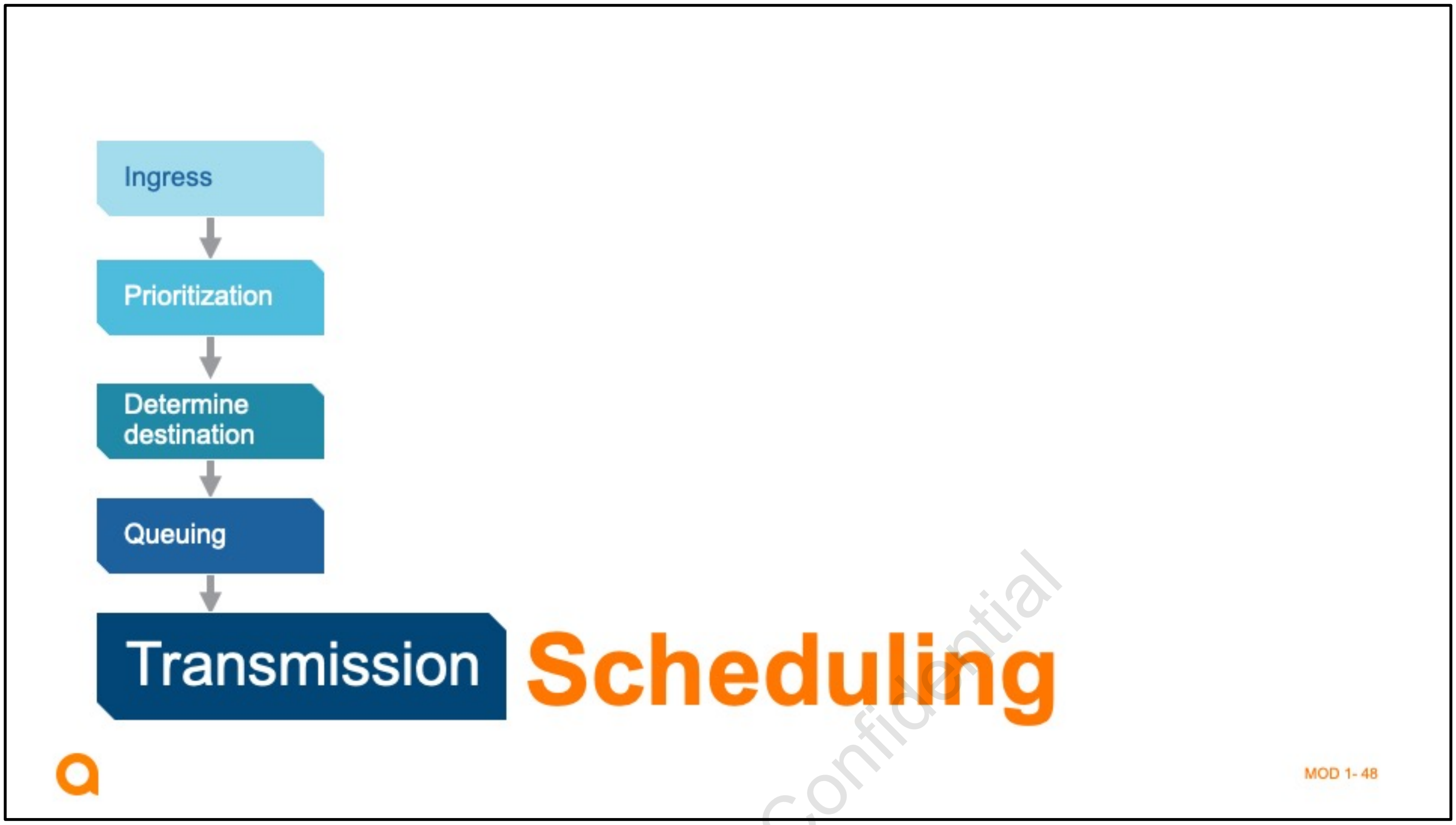
Schedule-Profile 1

Queue	Algorithm	Settings
0	WFQ	
1	WFQ	
…		
7		

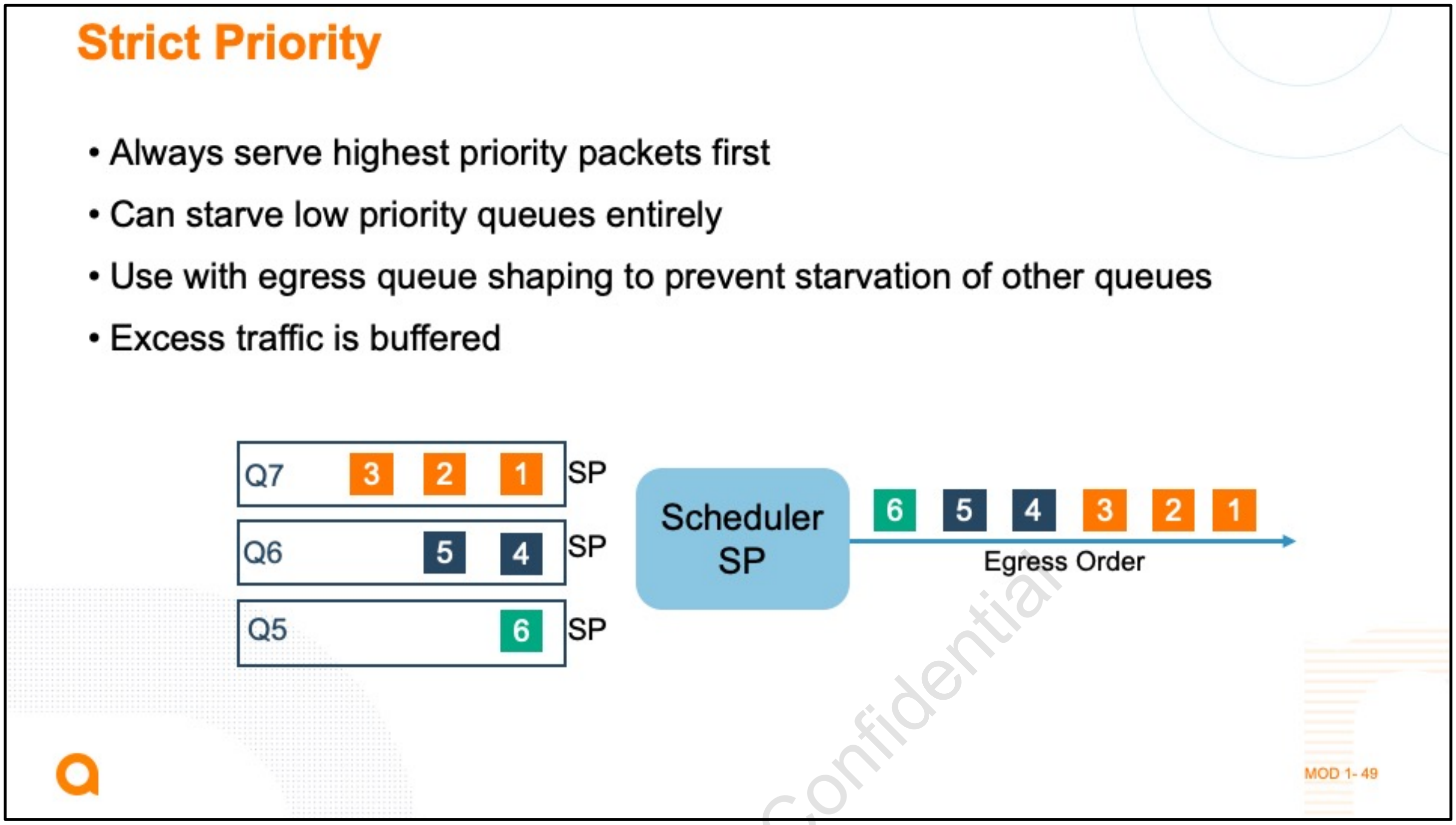
Now that traffic is in the correct queue, the port schedules it for forwarding. A port's schedule profile determines how the port determine the queue service order. You can apply a schedule profile to each individual port or LAG. LAG port members inherit the LAG schedule profile - you cannot apply a profile to LAG port members directly. If you do not apply a schedule profile to a port or LAG, it takes its schedule profile from the global profile.

The schedule profile must include an entry for every queue that you define in the global queue profile. The factory default profile has eight queues, numbered 0-7. For each queue, you specify a scheduling algorithm and settings associated with that algorithm. All the AOS-CX switch models support strict priority (SP) for the algorithm. In addition, the 8325, 8320, 6400, and 6300 support deficit weighted round robin (DWRR) while the 8400 supports weighted fair queuing (WFQ).

Every queue in the profile must use the same algorithm with one exception: the highest priority queue can use SP while the other queues use DWRR or WFQ. This is shown in the figure – a typical scenario. Some jitter/delay sensitive traffic like VoIP uses Strict Priority queuing, and the rest of the queues use either WFQ or DWRR. Basically, the Strict Priority Queue is serviced until it is empty. Then it uses WFQ or DWRR to service other queues, ensuring that all traffic types get their fair share of service.



Packets are queued, and so now they are scheduled to be transmitted.



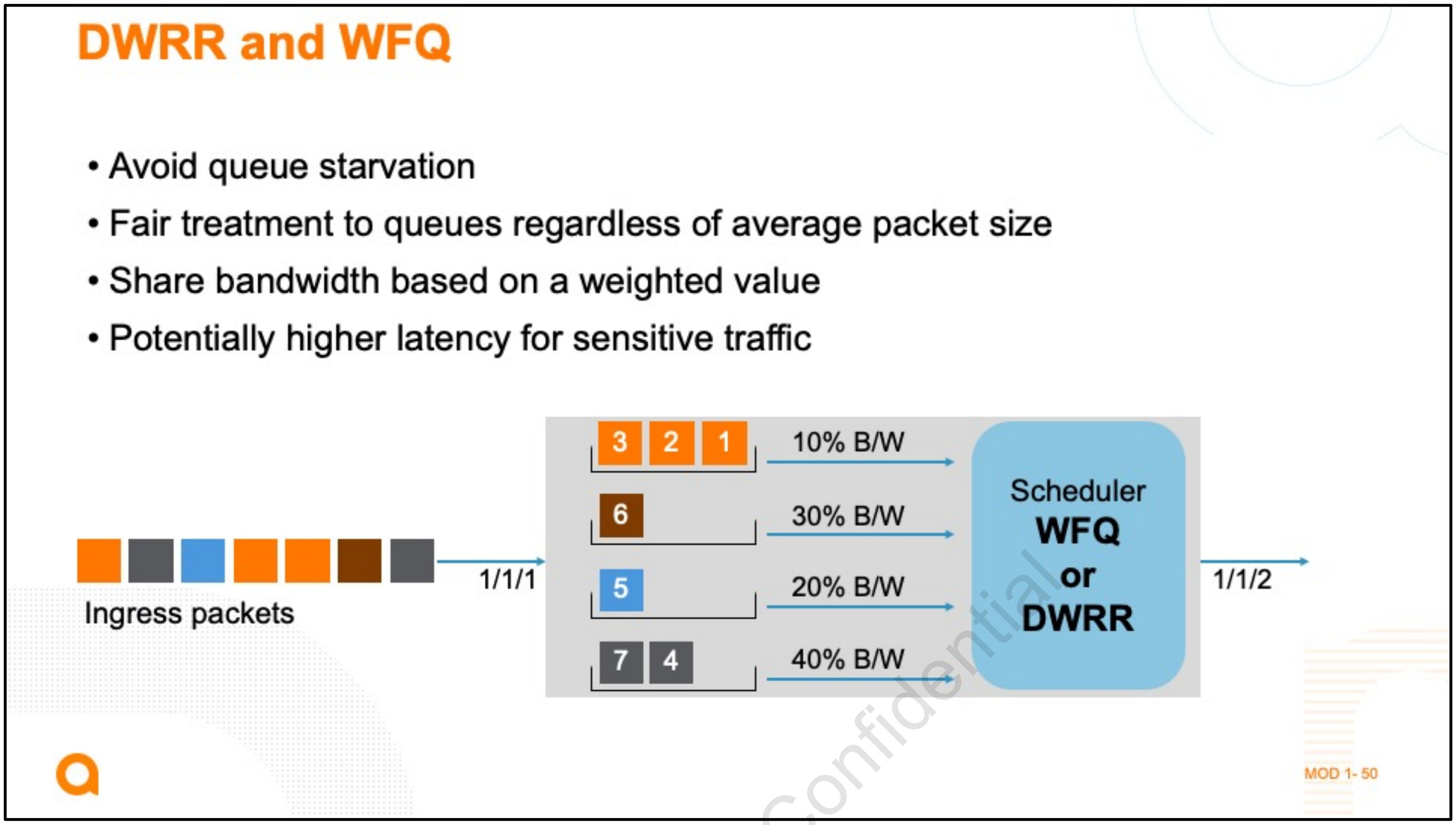
Strict Priority

- Always serve highest priority packets first
- Can starve low priority queues entirely
- Use with egress queue shaping to prevent starvation of other queues
- Excess traffic is buffered

SP scheduling always forwards the highest priority traffic first. That is, the port forwards all packets in queue 7. Packets Q6 are serviced if and only if Q7 is empty. Q5 is serviced if and only if packets in queues 7 and 6 are empty, and so on. If packets in Q3 are being serviced, and a packet arrives in Q7, the packet in the middle of being serviced is finished up, and then Q7 is serviced.

Under SP, the port will never forward a packet in a lower priority queue if packets show up in a higher priority queue. If queue 7 receives more packets before queue 4 has a chance to be served– the lower priority traffic is starved out entirely. Therefore, SP works best when the network assigns only a few applications, which take a relatively low amount of bandwidth, to the higher priority queues—network control traffic and voice over IP (VoIP). The switch then forwards this traffic immediately when it arrives, but still has bandwidth left over for other traffic in the best effort queue.

If you want to use SP alone, you could apply egress queue shaping to the high priority queues to prevent starvation of low priority queues. Egress queue shaping allows you to apply a maximum bandwidth to a priority queue, as well as a burst size. The port buffers excess traffic up to the burst size and sends the buffered traffic at the max rate, smoothing out bursts while also preventing the high priority queue from exceeding its maximum rate and starving out lower priority queues.



DWRR and WFQ

- Avoid queue starvation
- Fair treatment to queues regardless of average packet size
- Share bandwidth based on a weighted value
- Potentially higher latency for sensitive traffic

MOD 1- 50

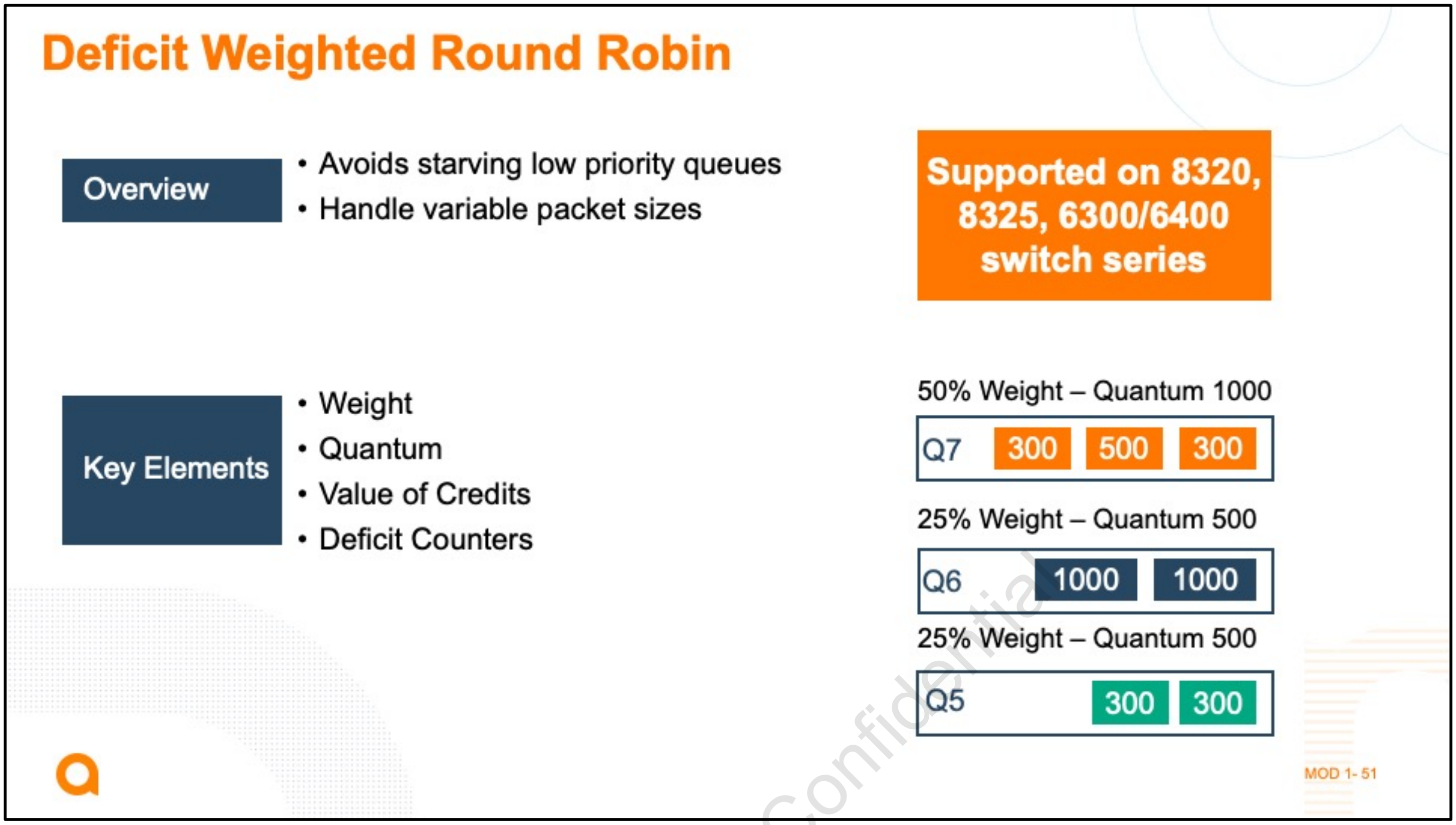
Deficit Weighted Round Robin (DWRR) and Weighted Fair Queuing (WFQ) both attempt to avoid the shortcomings of SP and ensure that even low priority queues receive a level of service. With both methods the scheduling algorithm work at the bit level, although in different ways. They aim to give fair treatment to every queue regardless of its average packet size. This principle is important for reducing jitter for small packets and can be crucial for applications such as VoIP.

In both algorithms, each queue receives a predictable share of the bandwidth based on the queue's relative priority, or weight. You can calculate the queue's bandwidth by taking the queue weight and dividing it by the total weight of all non-empty queues and then multiplying that by the port bandwidth.

Using weights that add up to 100 makes it easy to estimate the bandwidth: the weight converts to a percentage of the bandwidth. Pay attention to the "non-empty" part of the formula.

For example, a port has four queues with weights 40, 30, 20, and 10. If all queues are contending for bandwidth, the queue with weight 40 is guaranteed up to 40 percent of the bandwidth. If a queue empties and does not need its full share of the bandwidth, that bandwidth is available for other queues, shared out according to their weight.

Both DWRR and WFQ can impose a bit higher latency for high priority queues as compared to SP.



Deficit Weighted Round Robin

Overview

- Avoids starving low priority queues
- Handle variable packet sizes

Supported on 8320, 8325, 6300/6400 switch series

Key Elements

- Weight
- Quantum
- Value of Credits
- Deficit Counters

50% Weight – Quantum 1000

Q7 300 500 300

25% Weight – Quantum 500

Q6 1000 1000

25% Weight – Quantum 500

Q5 300 300

MOD 1- 51

Deficit algorithms can handle packets of variable size without knowing the mean size. A maximum packet size number is subtracted from the packet length, and packets that exceed that number are held back until the next scheduling turn.

Deficit Weighted Round Robin (DWRR) serves packets at the head of every non-empty queue whose deficit counter is greater than the size of the packet at the head of the queue. If the deficit counter is lower, the queue is skipped and its credit value, called a quantum, is increased. The increased value is used to calculate the deficit counter the next time around when the scheduler examines the queue for service. If the queue is served, the credit is decremented by the size of packet being served.

There are four key elements that affect how DWRR schedules queue service:

Weight. Reflects a proportion of the bandwidth on the outgoing interface.

Quantum. Translates the weight value into bytes. With each scheduling turn, a value of quantum is added in proportion to the weight. Thus, the quantum is the throttle mechanism that allows the scheduling to be aware of the bandwidth.

The value of credits can be positive or negative. Positive credits accrue when, at the end of a scheduler turn, there are leftover bytes that were not used in the queue’s scheduling turn. This value is deferred to the queue’s next scheduling turn. Negative credits accrue when the queue has transmitted more than its bandwidth value in a scheduling turn, and thus the queue is in debt when the next scheduling turn comes around.

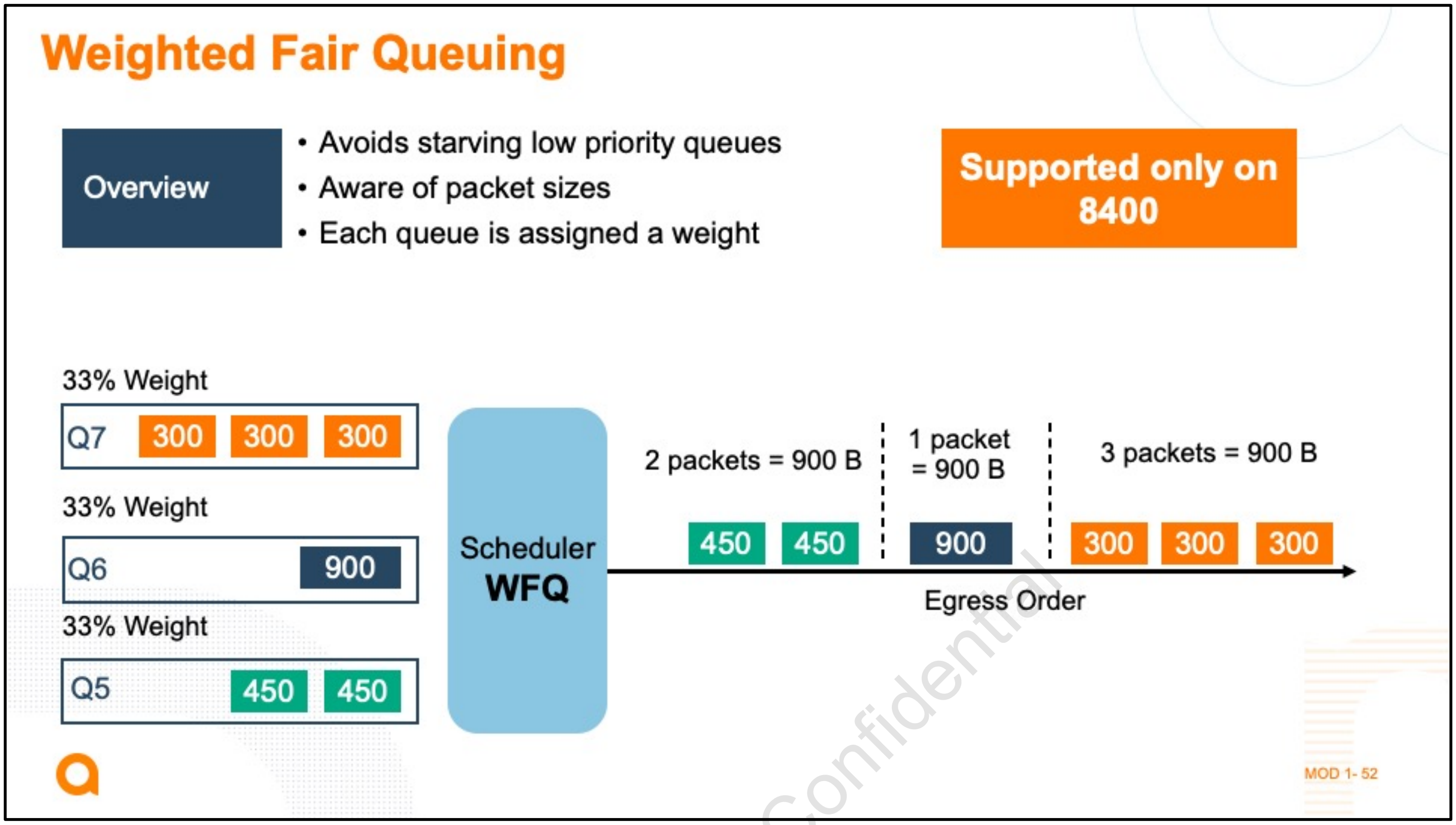
The deficit counter, which provides bandwidth fairness, is the sum of the quantum and the credits. The scheduler removes packets until the deficit counter reaches a value of zero or until the size of the packets is larger than the remaining deficits. But if the queue does not have enough deficits to schedule a packet, the value of the deficit is retained until the next scheduling round, and the scheduler moves to the next queue.

Benefits:

The DWRR algorithms limit the shortcomings with WRR because they are a more modern form of Weighted Fair Queuing (WFQ) that incorporates scheduling aware of bits and packets.

Limitation:

Services that have very strict demand on delay and jitter can be affected by other queues by the scheduling order. DWRR has no way to prioritize its scheduling.



Weighted Fair Queuing

Overview
- Avoids starving low priority queues
- Aware of packet sizes
- Each queue is assigned a weight

Supported only on 8400

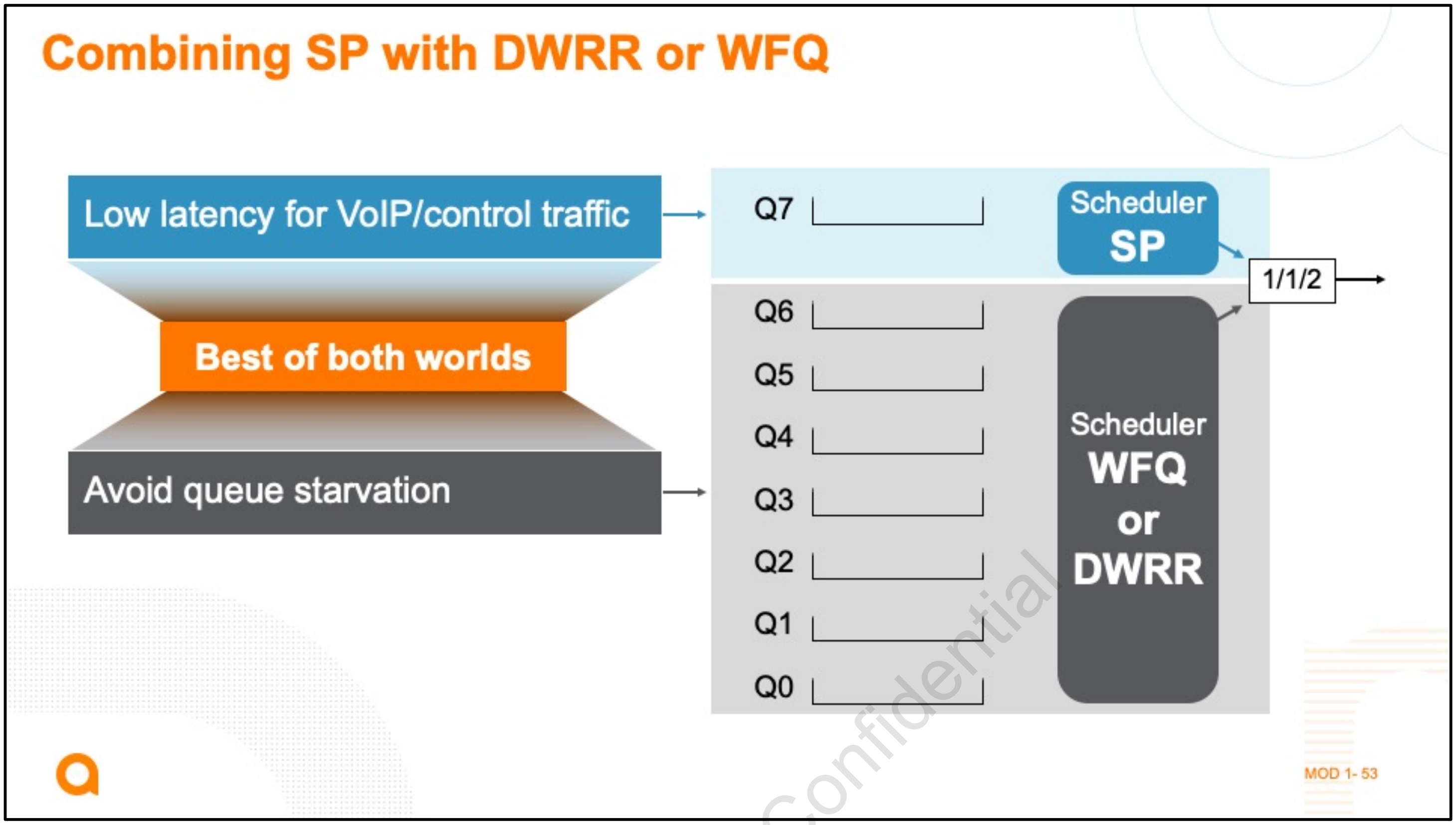
Weighted fair queuing (WFQ) is commonly referred as “bit-by-bit round robin,” because it implements a queuing and scheduling mechanism in which queue servicing is based on bits instead of packets.

In WFQ, each queue or flow is allocated a weight that is a proportion of the interface rate or the shaping rate. WFQ is aware of packet sizes and can thus support variable-sized packets. The benefit is that sessions with big packets do not get more scheduling time than sessions with smaller packets, because effectively the focus of WFQ is bits and not packets. So, there is no unfairness in the scheduling for sessions with smaller packet sizes. With WFQ, each queue is scheduled based on a computation performed on the bits of each packet at the head of the queue.

WFQ brings queue weight into the calculation of the virtual finish time. You can approximate how this process works by thinking of each queue as having its own bandwidth, which is its share of the port bandwidth relative to its weight. Based on this weighted bandwidth, queues with higher weight will have relatively sooner finish times for their packets than queues with lower weights.

In this example, all queues have the same weight, and thus the number of bytes scheduled in each scheduling turn is the same for all queues and reflects the weight value. Q7 removes three packets with a total value of 900 bytes, Q6 removes one packet with 900 bytes, and Q5 removes two packets with 900 bytes. The weight factor is effectively an allowance for how many resources can be assigned and used.

Note: Queuing does not fragment frames. If a queue could transmit 450 more bytes, but there is a 900-byte frame in the queue, it is not transmitted.



Combining SP with DWRR or WFQ

Low latency for VoIP/control traffic

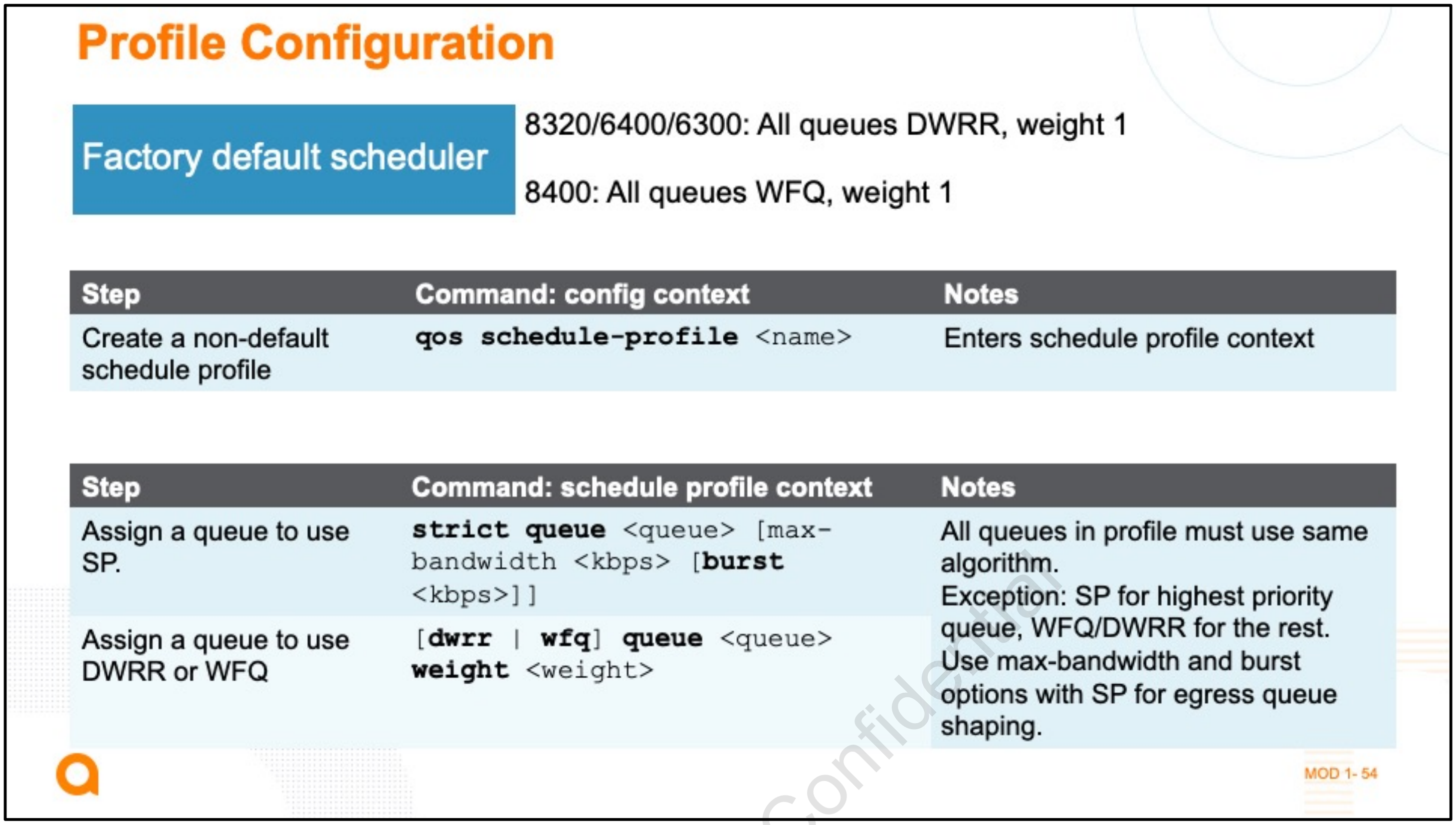
Best of both worlds

Avoid queue starvation

MOD 1- 53

For network has traffic that requires very low latency and jitter, such as VoIP traffic, you may want to use a Strict Priority (SP) queue for that traffic. This is because SP ensures high priority traffic the lowest latency and jitter. However, you may not want to use SP for medium high priority queues due to the risk of starving out other traffic.

AOS-CX switches give you the best of both worlds. You can create a schedule profile that combines either DWRR or WFQ, depending on the switch model, with Strict Priority. In a hybrid profile, only the highest priority queue can use SP. This is queue 7 if the profile uses all queues or the highest queue used if the profile does not. You should assign traffic like VoIP to this queue; network control traffic is also automatically assigned to it. The switch forwards traffic from this queue as long as the queue has traffic. When the queue is empty, the port uses the remaining bandwidth to serve other queues. It allocates this remaining bandwidth using DWRR or WFQ. Because VoIP and network control traffic typically have low bandwidth requirements, the risks of starving the other traffic is minimal.



Profile Configuration

Factory default scheduler

8320/6400/6300: All queues DWRR, weight 1

8400: All queues WFQ, weight 1

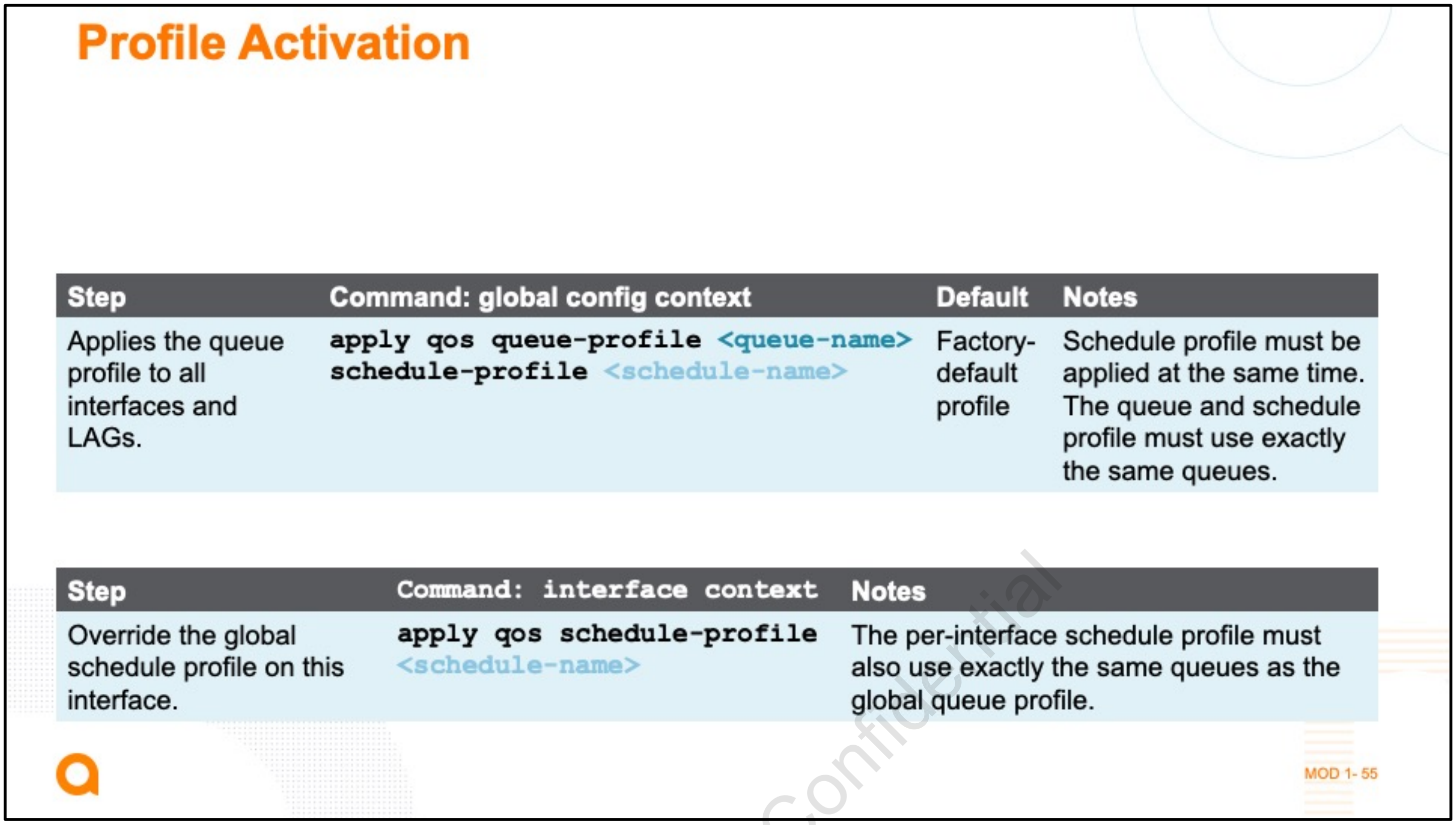
Step	Command: config context	Notes
Create a non-default schedule profile	`qos schedule-profile <name>`	Enters schedule profile context

Step	Command: schedule profile context	Notes
Assign a queue to use SP.	`strict queue <queue> [max-bandwidth <kbps> [burst <kbps>]]`	All queues in profile must use same algorithm. Exception: SP for highest priority queue, WFQ/DWRR for the rest. Use max-bandwidth and burst options with SP for egress queue shaping.
Assign a queue to use DWRR or WFQ	`[dwrr | wfq] queue <queue> weight <weight>`	

MOD 1- 54

The 832X, 6400, and 6300 AOS-CX switches default scheduling profile uses DWRR for all queues with all DWRR queue weights the same. The 8400 AOS-CX switches default scheduling profile uses WFQ with all queue weights the same. Clearly you will need to create your own profile or profiles to give the higher priority queues the preference that you desire.

Note: Currently, only the AOS-CX 8400 switches support WFQ; the other AOS-CX switches support DWRR.



Profile Activation

Step	Command: global config context	Default	Notes
Applies the queue profile to all interfaces and LAGs.	`apply qos queue-profile <queue-name> schedule-profile <schedule-name>`	Factory-default profile	Schedule profile must be applied at the same time. The queue and schedule profile must use exactly the same queues.

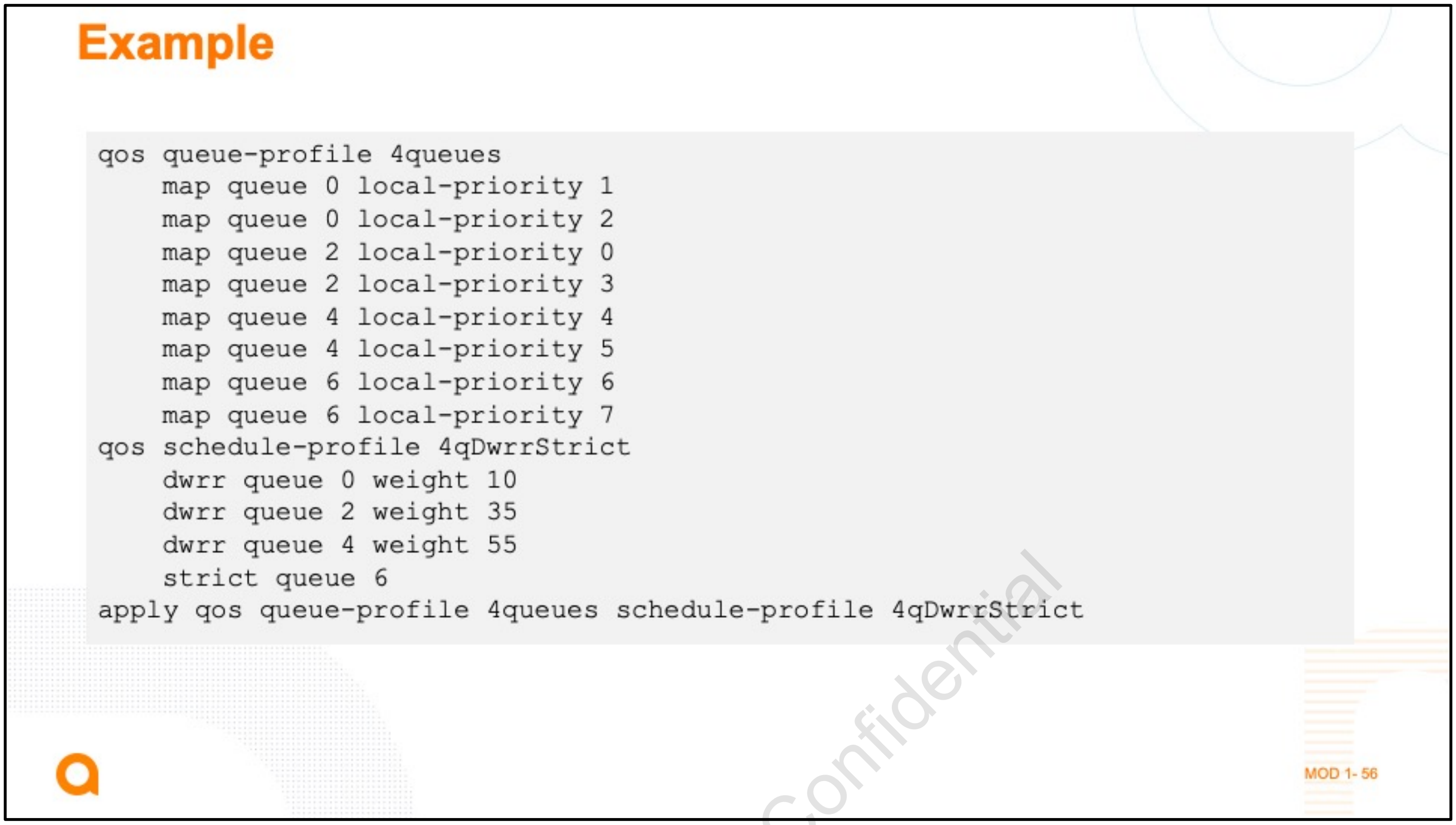
Step	Command: interface context	Notes
Override the global schedule profile on this interface.	`apply qos schedule-profile <schedule-name>`	The per-interface schedule profile must also use exactly the same queues as the global queue profile.

MOD 1- 55

You'll apply the global schedule profile when you apply the queue profile. You need to apply the queue and schedule profile together. The queue and schedule profile must use exactly the same queues.

If you want, you can keep the factory-default queue profile and apply a new schedule profile with this command: apply qos queue-profile factory-default schedule-profile <your profile name>.

If you want to override the global schedule profile on a particular physical interface or LAG, access its context and enter the apply qos schedule-profile command. Remember that the global schedule profile and individual interface schedule profiles must use exactly the same queues that you used in the queue profile. For the factory-default queue profile, there are eight queues from 0 to 7.

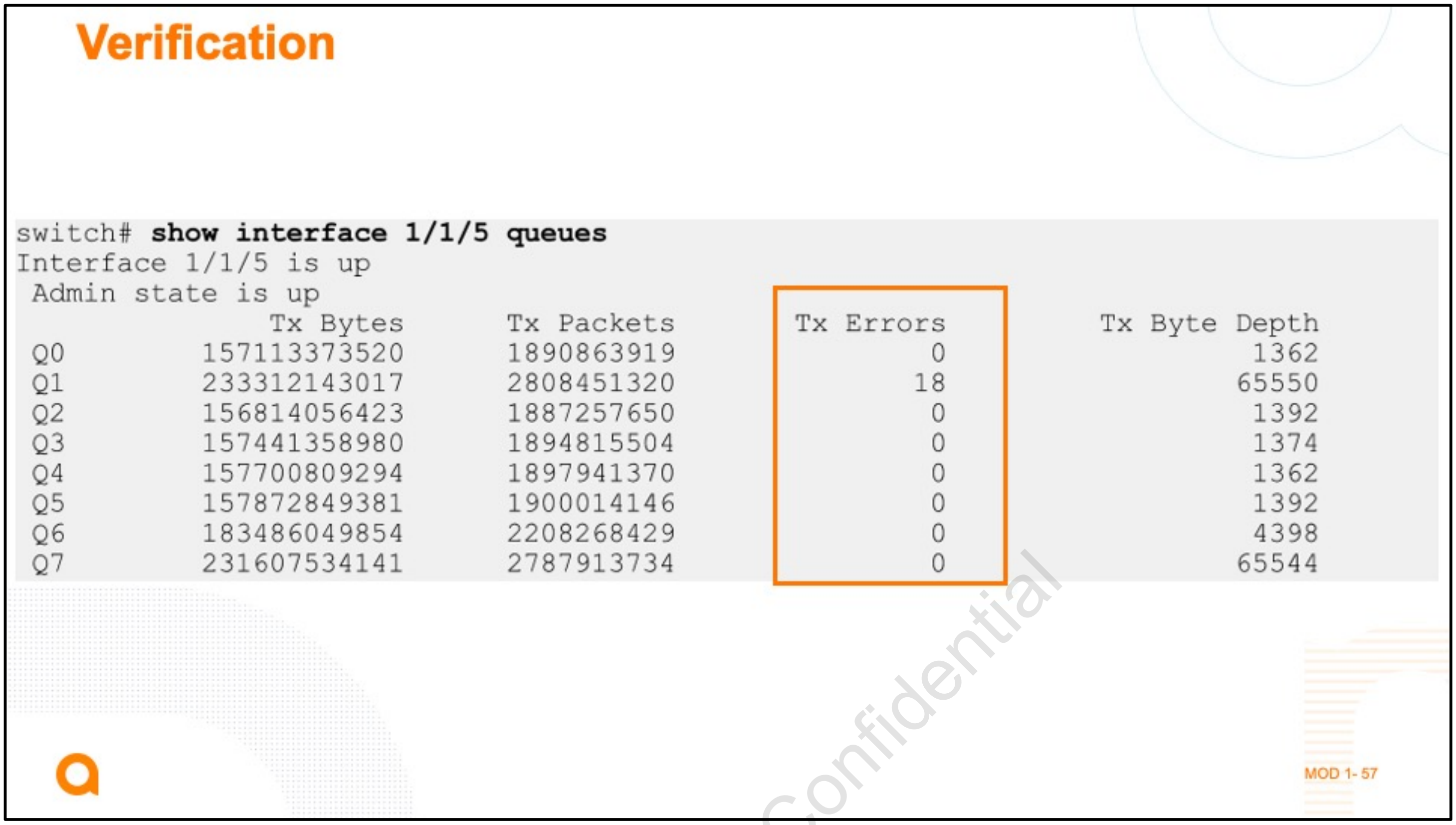


Example

```
qos queue-profile 4queues
    map queue 0 local-priority 1
    map queue 0 local-priority 2
    map queue 2 local-priority 0
    map queue 2 local-priority 3
    map queue 4 local-priority 4
    map queue 4 local-priority 5
    map queue 6 local-priority 6
    map queue 6 local-priority 7
qos schedule-profile 4qDwrrStrict
    dwrr queue 0 weight 10
    dwrr queue 2 weight 35
    dwrr queue 4 weight 55
    strict queue 6
apply qos queue-profile 4queues schedule-profile 4qDwrrStrict
```

MOD 1- 56

Here is an example configuration for a custom queue and schedule profile that works with that queue profile. Notice that only four (4) queues are being used (0, 2, 4, and 6). Based on the schedule profile, DWRR is being used and the queue and schedule profile are applied globally.



Verification

```
switch# show interface 1/1/5 queues
Interface 1/1/5 is up
 Admin state is up
               Tx Bytes        Tx Packets        Tx Errors            Tx Byte Depth
 Q0         157113373520      1890863919                0                     1362
 Q1         233312143017      2808451320               18                    65550
 Q2         156814056423      1887257650                0                     1392
 Q3         157441358980      1894815504                0                     1374
 Q4         157700809294      1897941370                0                     1362
 Q5         157872849381      1900014146                0                     1392
 Q6         183486049854      2208268429                0                     4398
 Q7         231607534141      2787913734                0                    65544
```

MOD 1- 57

Use the show interface command to display the number of packets and bytes queued for transmission.

Here’s a description of each of the column output for the queues:

Tx Bytes: Total bytes transmitted. The byte count may include packet headers and internal metadata that are removed before the packet is transmitted. Packet headers added when the packet is transmitted may not be included.

Tx Packets: Total packets transmitted.

Tx Errors: Shows the amount of traffic dropped on an egress interface before it was sent. When traffic cannot be forwarded out an egress interface, it backs up on ingress. The more servicing assigned to a queue by a schedule profile, the less likely traffic destined for that queue will back up and be dropped. Tx Errors shows the sum of packets that were dropped across all line modules (due to insufficient capacity) by the ingress Virtual Output Queues (VOQs) destined for the egress port. As the counts are read separately from each line module, the sum is not an instantaneous snapshot.

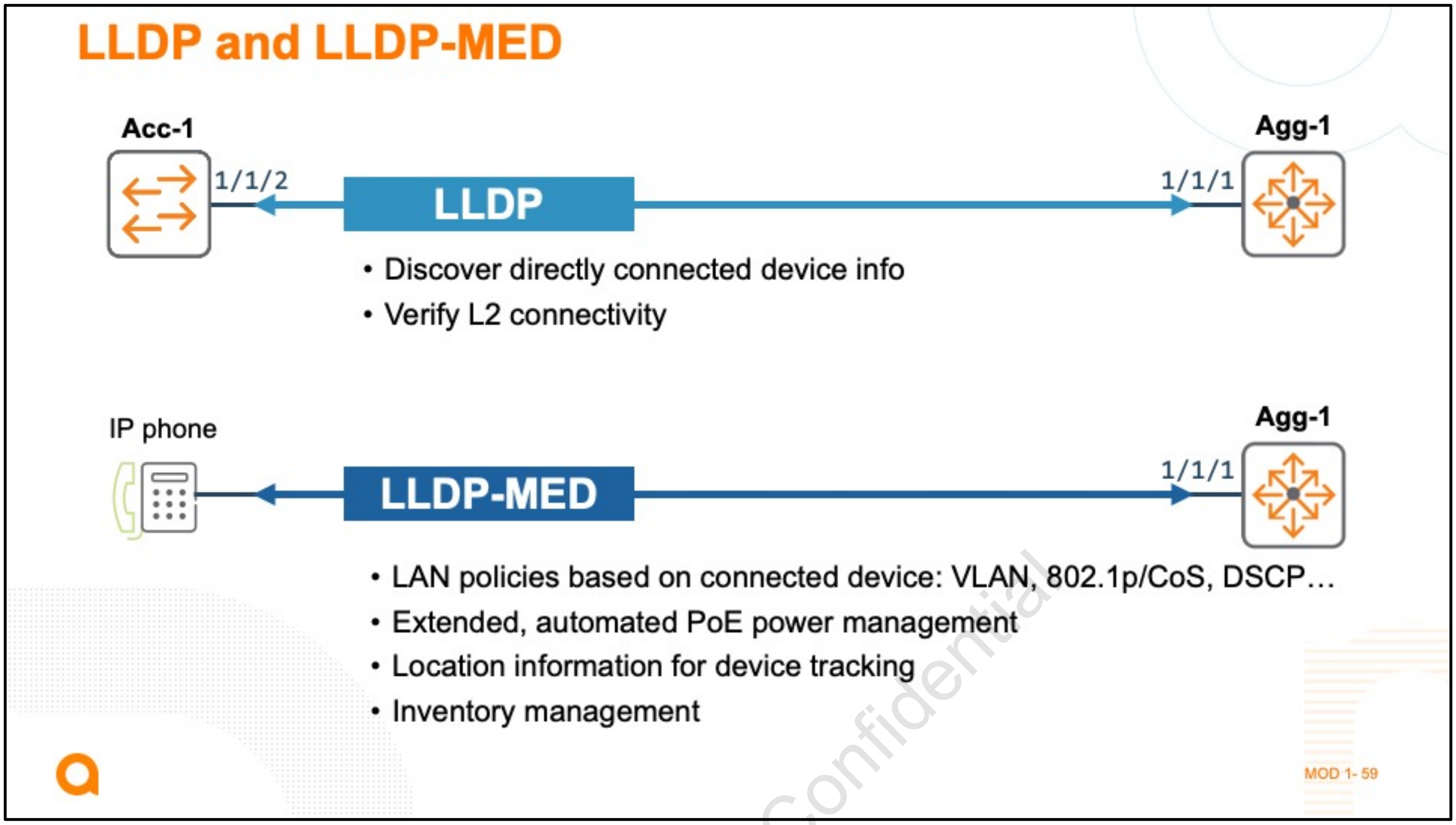
Tx Byte Depth: Largest byte depth (or high watermark) found on any ingress line module VOQ destined for the egress port.

.

LLDP-MED and Device Profiles

MOD 1- 58

You have learned how to classify, prioritize, and apply policies to traffic. You will now learn how to support Voice over IP (VoIP) phones using LLDP-MED.



LLDP and LLDP-MED

- Discover directly connected device info
- Verify L2 connectivity

- LAN policies based on connected device: VLAN, 802.1p/CoS, DSCP…
- Extended, automated PoE power management
- Location information for device tracking
- Inventory management

MOD 1- 59

The Link Layer Discovery Protocol (LLDP) is an open-standard link layer protocol used by network devices for advertising their identity, capabilities, and neighbors on a local area network based on IEEE 802 technology, primarily Ethernet. It can be used to verify Layer-2 connectivity between devices as well as learning basic information about directly-connected devices.

The Media Endpoint Discovery (MED) is an enhancement of LLDP, known as LLDP-MED. It is commonly used to provision devices, like determine the correct power settings for a connected VoIP phone or AP, or automatically assigning the correct VLAN and QoS settings for a connected device. It can enable plug-and-play networking. Features include the following:

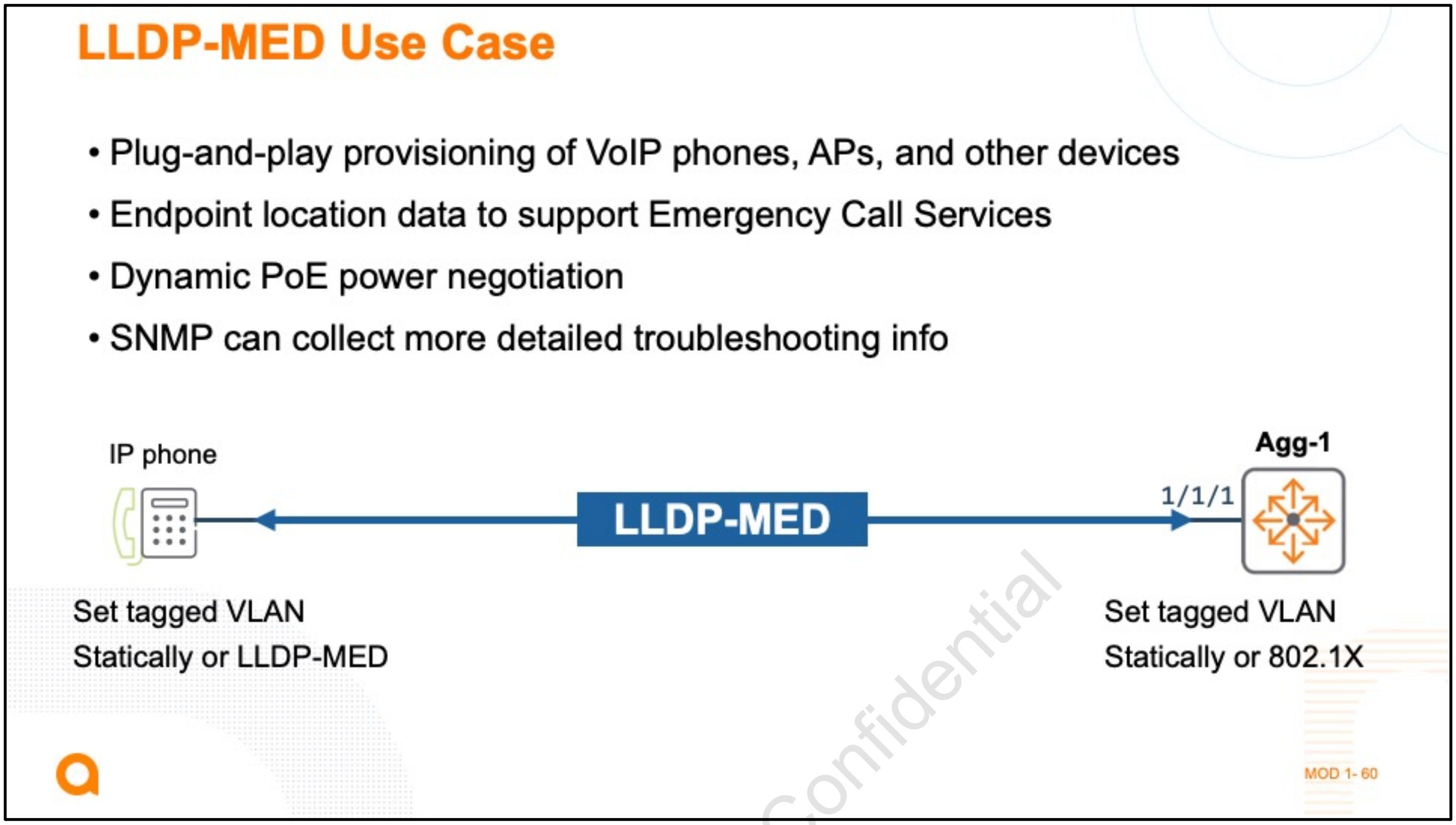
Set LAN policies based on the connected device: VLAN, Layer 2 802.1p CoS values, Layer 3 DSCP settings, and more.

Device location discovery allows the creation of location databases: VoIP and Enhanced 911 services.

Extended and automated power management of Power-over-Ethernet (PoE) end devices.

Inventory management: Track your network devices, and determine device-specific info - manufacturer, software and hardware versions, serial or asset number.

Support for the fast-start capability



LLDP-MED Use Case

- Plug-and-play provisioning of VoIP phones, APs, and other devices
- Endpoint location data to support Emergency Call Services
- Dynamic PoE power negotiation
- SNMP can collect more detailed troubleshooting info

Set tagged VLAN
Statically or LLDP-MED

Set tagged VLAN
Statically or 802.1X

MOD 1- 60

You know to place VoIP phones in their own VLAN and configure a high priority for that VLAN – typically a tagged VLAN. You connect a PC via Ethernet cable to the phone. The PC sends its frames on a native, untagged VLAN. But setting up the switch port to receive the tagged traffic is not enough. The phone must also know what VLAN ID to use for its packets. You can manually configure that on each phone – tedious.

With LLDP-MED, the switch automatically sends the correct VLAN ID to the IP phone, along with 802.1p and DSCP settings. You get automatic, plug-and-play provisioning for LLDP-MED capable VoIP phones. The switch can also communicate other information to the phone, including its physical location. This information can help to support Emergency Call Services (ECS).

When phones receive Power over Ethernet (PoE) from the switch, LLDP-MED can help the switch allocate and deliver exactly the power that the phone needs.

LLDP-MED can also help in tracking and troubleshooting VoIP phones through a central management platform. SNMP servers can read detailed VoIP endpoint data inventory, PoE status, and information to help troubleshoot the IP telephony network and call quality issues.

Review interoperation with authentication

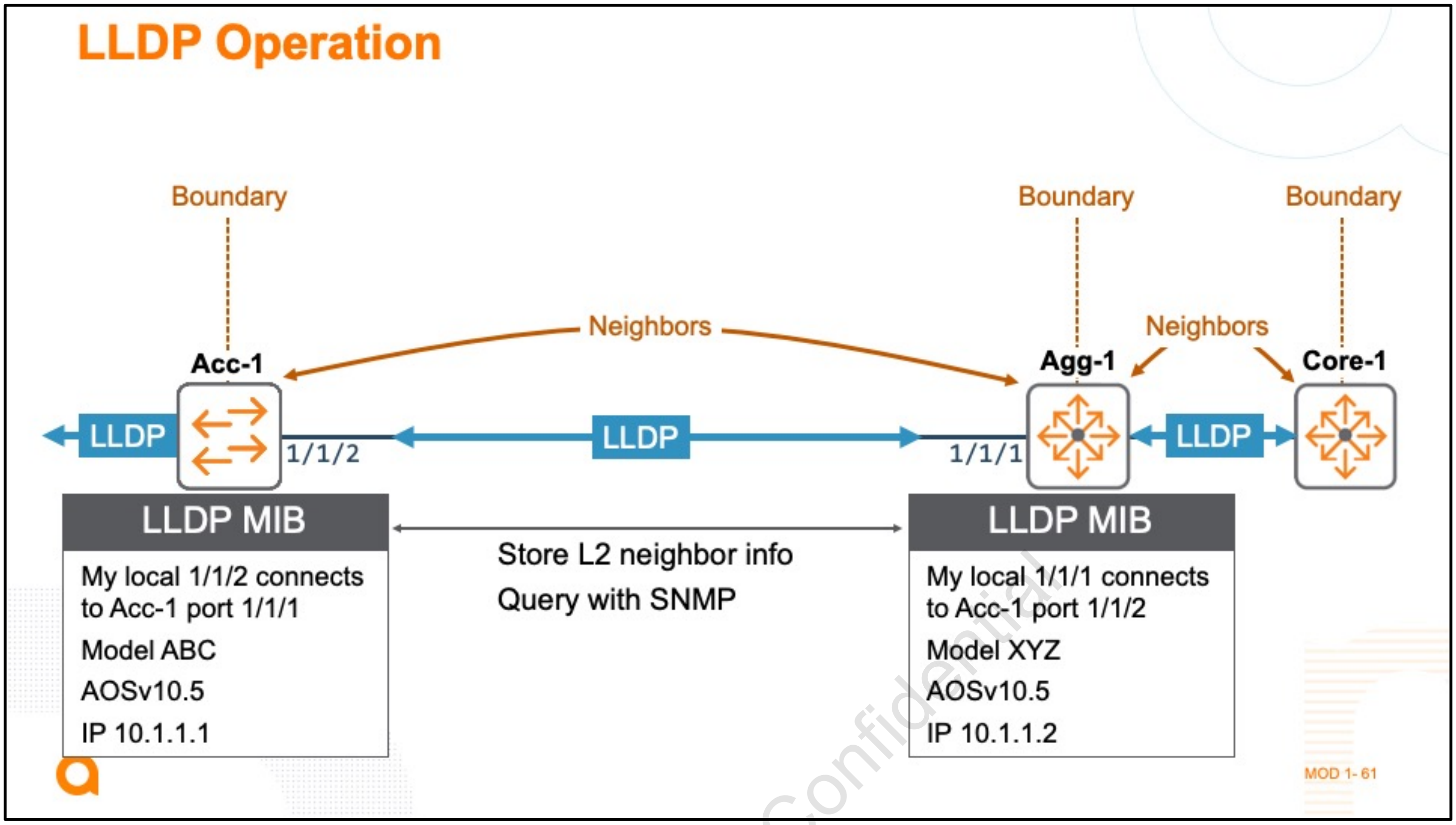
You manually configure switch ports with the tagged VLAN for VoIP, or rely on dynamic assignment using 8021.X and RADIUS attributes.

You manually configure each IP phone with a matching tagged VLAN, or rely on LLDP-MED to automatically configure it for you.

You can use 802.1X with LLDP-MED. Simply set up 802.1X on the port. Make sure that IP phones support 802.1X and that the RADIUS server has policies and accounts for authenticating them. If you want the server to send the tagged VLAN assignment dynamically, make sure that it has the proper policies for doing so.

For this scenario, you probably want the computer and IP phone to authenticate separately so that an unauthorized user cannot piggyback on the IP phone's session. Make sure to set the 802.1X client-limit to 2 so that the port operates in user-mode and authenticates each device separately.

If you want to enforce 802.1X but the IP phones do not support it, combine 802.1X for users and MAC-Auth for IP phones.



LLDP Operation

LLDP MIB

My local 1/1/2 connects to Acc-1 port 1/1/1

Model ABC

AOSv10.5

IP 10.1.1.1

Store L2 neighbor info

Query with SNMP

LLDP MIB

My local 1/1/1 connects to Acc-1 port 1/1/2

Model XYZ

AOSv10.5

IP 10.1.1.2

The IEEE 802.1AB Link Layer Discovery Protocol (LLDP) provides a standards-based method for network devices to discover each other and exchange information about their capabilities. Devices send LLDP data packets out all interfaces with outbound LLDP enabled. Devices accept LLDP advertisements from L2-connected neighbors on ports with inbound LLDP enabled.

Inbound packets from neighbor devices are stored in a special LLDP Management Information Base (MIB). This information can then be queried by other devices through SNMP. Network management tools can use LLDP information to create accurate physical network topologies. LLDEP helps to determine which devices are neighbors and through which interfaces they connect. LLDP operates at layer 2 and requires an LLDP agent to be active on each interface that sends and receives LLDP advertisements.

LLDP advertisements can contain a variable number of TLV (type, length, value) information elements. Each TLV describes a single attribute of a device such as: system capabilities, management IP address, device ID, port ID.

Packet boundaries

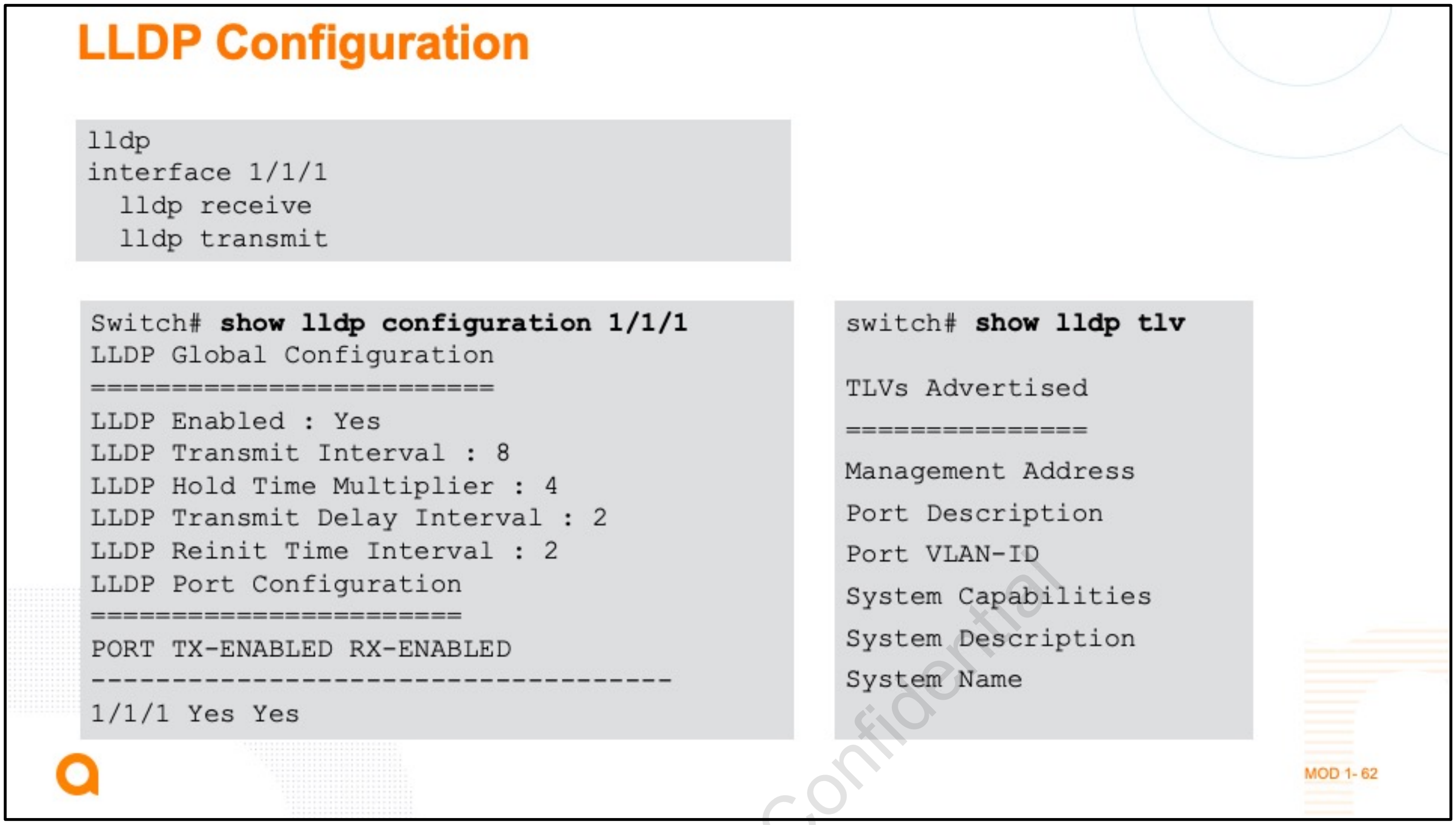
LLDP packets travel only to directly-connected peers. These packets are never forwarded to any other devices, regardless of whether they are LLDP-enabled.

An intervening hub or repeater forwards the LLDP packets it receives in the same manner as any other multicast packets it receives. Therefore, two LLDP switches joined by a hub or repeater handle LLDP traffic in the same way that they would if directly connected.

Any intervening 802.1D device or Layer-3 device that is either LLDP-unaware or has disabled LLDP operation drops the packet.

LLDP-MED

LLDP-MED in the switches uses the standard LLDP commands described earlier in this section, with some extensions, and also introduces new commands unique to LLDP-MED operation.



LLDP Configuration

```
lldp
interface 1/1/1
  lldp receive
  lldp transmit
```

```
Switch# show lldp configuration 1/1/1
LLDP Global Configuration
=========================
LLDP Enabled : Yes
LLDP Transmit Interval : 8
LLDP Hold Time Multiplier : 4
LLDP Transmit Delay Interval : 2
LLDP Reinit Time Interval : 2
LLDP Port Configuration
=======================
PORT TX-ENABLED RX-ENABLED
------------------------------------
1/1/1 Yes Yes
```

```
switch# show lldp tlv

TLVs Advertised
===============
Management Address
Port Description
Port VLAN-ID
System Capabilities
System Description
System Name
```

MOD 1- 62

LLDP is enabled by default, both globally and on interfaces. Configure LLDP if needed using the syntax shown in the figure. Validate as shown in the figure.

A portion of the output is not shown. At the bottom of the show LLDP configuration you would see the LLDP advertised TLVs. But you can also use the show lldp tlv command, as shown in the figure. It is good that the switch advertises information about itself, but it is not currently advertising TLVs for LLDP-MED

You should look for the following TLVs:

MED Capability: The capabilities TLV enables the switch to detect LLDP-MED capable endpoints, their class, and their capabilities.

MED Network Policy: The switch communicates the proper VLAN ID and QoS settings in this TLV. You will learn how to configure those on the next page.

MED Power (PoE): The switch advertises its capabilities as a power sourcing equipment (PSE). The VoIP phone can also use this TLV to request a specific amount of power after it first powers up or dynamically throughout the connection. Note that PoE+ defines its own LLDP TLV for negotiating the power (a dot3 TLV called poeplus_config), and the phone might use this TLV instead. In either case, the switch will allocate the requested amount of power and reserve it for this port even if the phone is not drawing the full power at the moment.

If you need to enable any of these TLVs, use these commands:

```
Switch(config)# interface <interface-ID>
Switch(config-if)# lldp med capability
Switch(config-if)# lldp med network-policy
Switch(config-if)# lldp med poe [priority-overide]
```

capability specifies advertisement of supported LLDP MED TLVs. The capability TLV is always sent with other MED TLVs, therefore it cannot be disabled when other MED TLVs are enabled. Default: enabled.

network-policy lets endpoints and network devices advertise their VLAN IDs, and IEEE 802.1p (PCP and DSCP) values for voice applications. This TLV is only sent when a voice VLAN policy is present. It is enabled, by default.

The priority-override option overrides user-configured port priority for Power over Ethernet. When both lldp dot3 poe and lldp med poe are enabled, the lldp dot3 poe3 setting takes precedence.

Use the lldp select-tlv command to change what TLVs the switch will send and receive. See the Fundamentals Guide for more information on this command. The show lldp tlv command shows the LLDP TLVs that are configured for send and receive:

```
Switch# show lldp tlv
TLVs Advertised
===============
Management Address
Port Description
Port VLAN-ID
System Capabilities
System Description
System Name
OUI
```

MED location TLV

It is best practice to populate the location ID TLV for each switch port that might connect to a VoIP phone with the correct location information. For example, the command for specifying the civic address of a media endpoint such as a phone is:

```
Switch(config)# interface <interface-ID>
Switch(config-if)# lldp med-location civic-addr <country-code> <CA type> <CA value> [<CA type> <CA value>]
```

A CA type specifies the type of location being indicated, such as 3 for a city. You can specify many CA types and values in order to program the city, street, address, and building, for example. For a list of CA types, see your switch Management and Configuration Guide.

You can alternatively specify an Emergency Line Identification Number (ELIN), which is used in North America, if such numbers have been assigned to your phones:

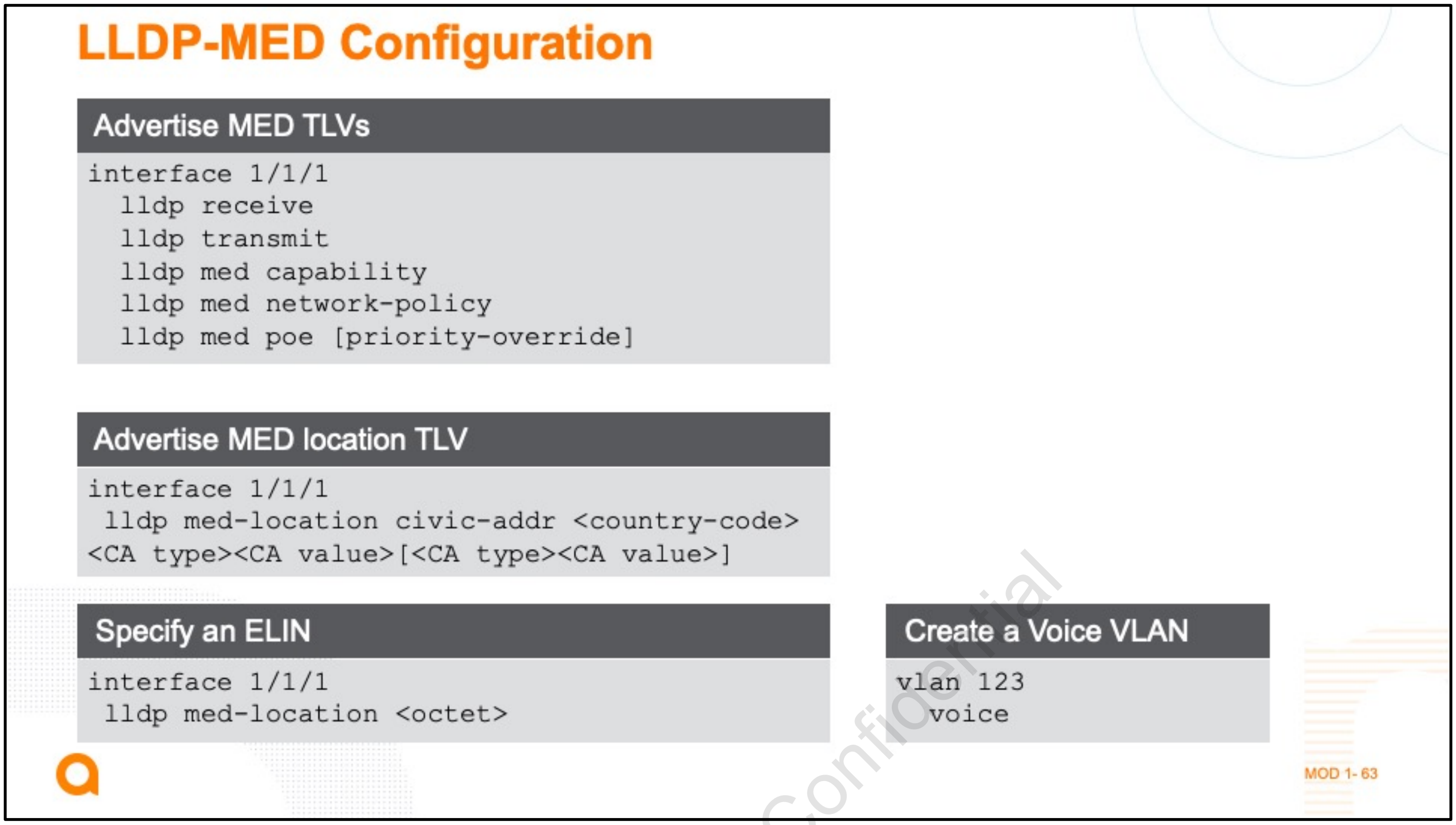
```
Switch(config)# interface <interface-ID>
Switch(config-if)# lldp med-location <octet>
```

Voice VLANs

To create a voice VLAN, configure the voice command in the VLAN context, like this:

```
Switch(config)# vlan <VLAN-ID>
Switch(config-vlan)# voice
```

LLDP-MED extensions support voice VLANs.



LLDP-MED Configuration

Advertise MED TLVs

```
interface 1/1/1
  lldp receive
  lldp transmit
  lldp med capability
  lldp med network-policy
  lldp med poe [priority-override]
```

Advertise MED location TLV

```
interface 1/1/1
 lldp med-location civic-addr <country-code>
<CA type><CA value>[<CA type><CA value>]
```

Specify an ELIN

```
interface 1/1/1
 lldp med-location <octet>
```

Create a Voice VLAN

```
vlan 123
  voice
```

You should look for the following TLVs related to LLDP-MED:

MED Capability: Enables the switch to detect LLDP-MED capable endpoints, their class, and their capabilities.

MED Network Policy: The switch send VLAN ID and QoS settings in this TLV.

MED Power (PoE): The switch advertises its capabilities as a Power Sourcing Equipment (PSE). The VoIP phone can also use this TLV to request a specific amount of power after it first powers up or dynamically throughout the connection. Note that PoE+ defines its own LLDP TLV for negotiating the power (a dot3 TLV called poeplus_config), and the phone might use this TLV instead. In either case, the switch will allocate the requested amount of power and reserve it for this port even if the phone is not drawing the full power at the moment.

To enable these TLVs, use the configuration shown in the figure.

capability – advertise supported LLDP MED TLVs. The capability TLV is always sent with other MED TLVs, therefore it cannot be disabled when other MED TLVs are enabled. Default: enabled.

network-policy lets endpoints and network devices advertise their VLAN IDs, and IEEE 802.1p (PCP and DSCP) values for voice applications. This TLV is only sent when a voice VLAN policy is present. It is enabled, by default.

The priority-override option overrides user-configured port priority for Power over Ethernet.

When both lldp dot3 poe and lldp med poe are enabled, the lldp dot3 poe3 setting takes precedence.

Use the lldp select-tlv command to change what TLVs the switch will send and receive. See the Fundamentals Guide for more information on this command.

MED location TLV

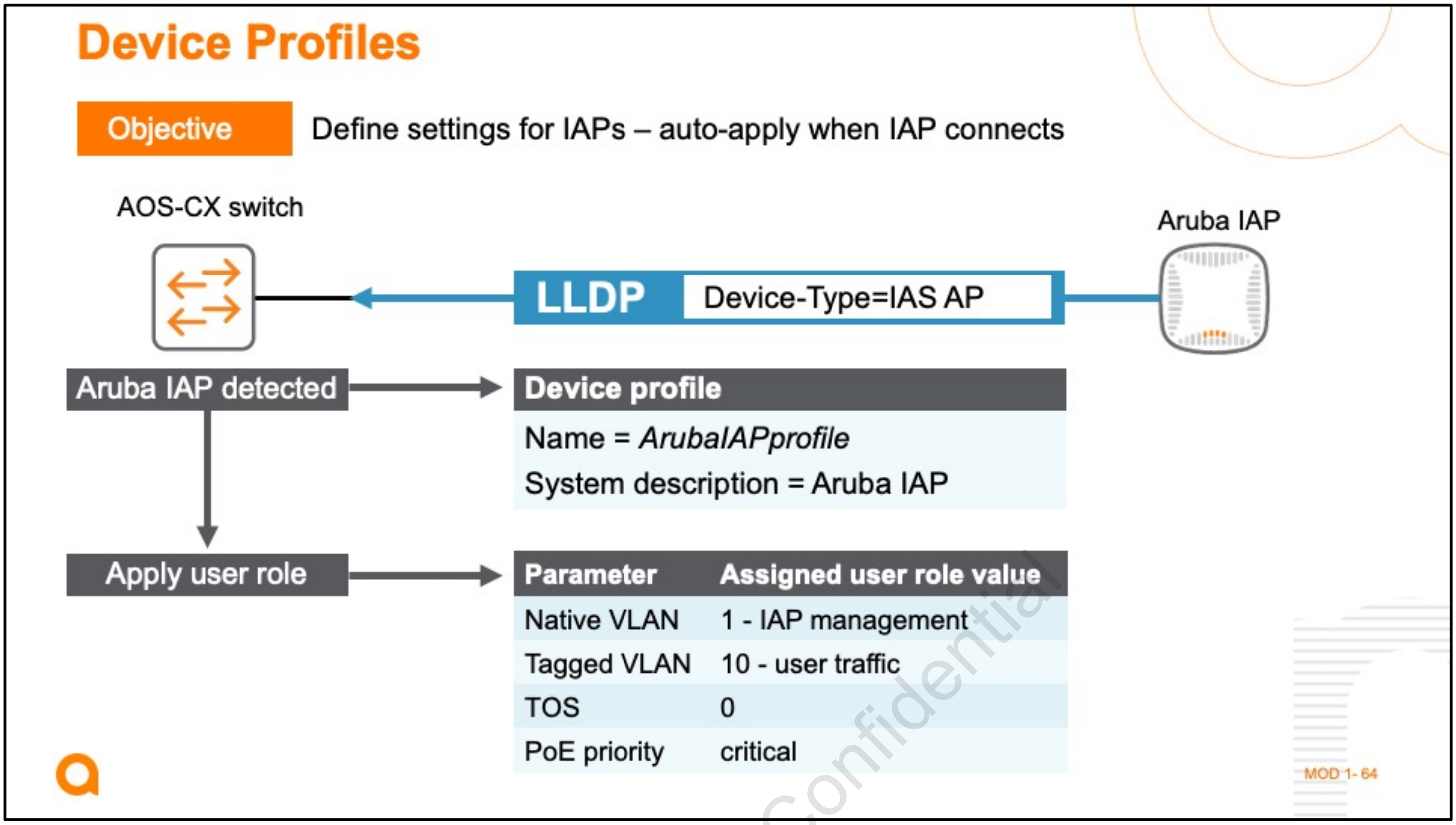
It is best practice to populate the location ID TLV for each switch port that might connect to a VoIP phone with the correct location information. For example, the command for specifying the civic address of a media endpoint such as a phone is shown in the figure.

A CA type specifies the type of location being indicated, such as 3 for a city. You can specify many CA types and values in order to program the city, street, address, and building, for example. For a list of CA types, see your switch Management and Configuration Guide.

You can alternatively specify an Emergency Line Identification Number (ELIN), which is used in North America, if such numbers have been assigned to your phones, as shown in the figure.

Voice VLANs

To create a voice VLAN, configure the voice command in the VLAN context, as shown in the figure. LLDP-MED extensions support voice VLANs.



Device Profiles

Objective Define settings for IAPs – auto-apply when IAP connects

Device profile

Name = *ArubaIAPprofile*

System description = Aruba IAP

Parameter	Assigned user role value
Native VLAN	1 - IAP management
Tagged VLAN	10 - user traffic
TOS	0
PoE priority	critical

MOD 1- 64

AOS-CX supports device profiles to make it even simpler to deploy Aruba Instant APs (IAP) and other devices. Use it when you are not sure what switch port the device might connect to. Typically, you have a standard configuration that applies to all AP-connected ports. This would include the native, untagged VLAN where IAPs have their IP addresses. It also includes tagged VLAN assignments for static and dynamic VLANs assigned to WLANs for Aruba IAPs, the PoE settings such as a critical PoE priority, and so on.

A device profile lets you define these settings once on a switch, before connecting the APs. The switch knows to apply settings when it detects a target device connect. You can match on a variety of criteria, but the most common is the LLDP device type field. You associate the device profile with the Aruba IAPs system description, for example, in its LLDP message. Installers can then connect IAPs to any switch port without worry. The switch detects the Aruba IAP type in an AP's LLDP messages and automatically applies the settings in the device profile (i.e., the user role associated with the device profile).

The switch provides support for LLDP and CDP to enable automatic discovery and configuration of other devices on the network.

When the LLDP information on the interface ages out, the profile/role is revoked from the interface. Only devices directly connected to the switch are detected and processed for device profile application.

You can apply the following interface settings when a target device connects:

Native (untagged) VLANs: The VLAN on which untagged frames (normal Ethernet frames) are classified. The process for applying the setting depends on how the VLAN interface is configured. See VLAN interfaces.

Tagged VLANs: The list of allowed tagged VLANs on the interface.*

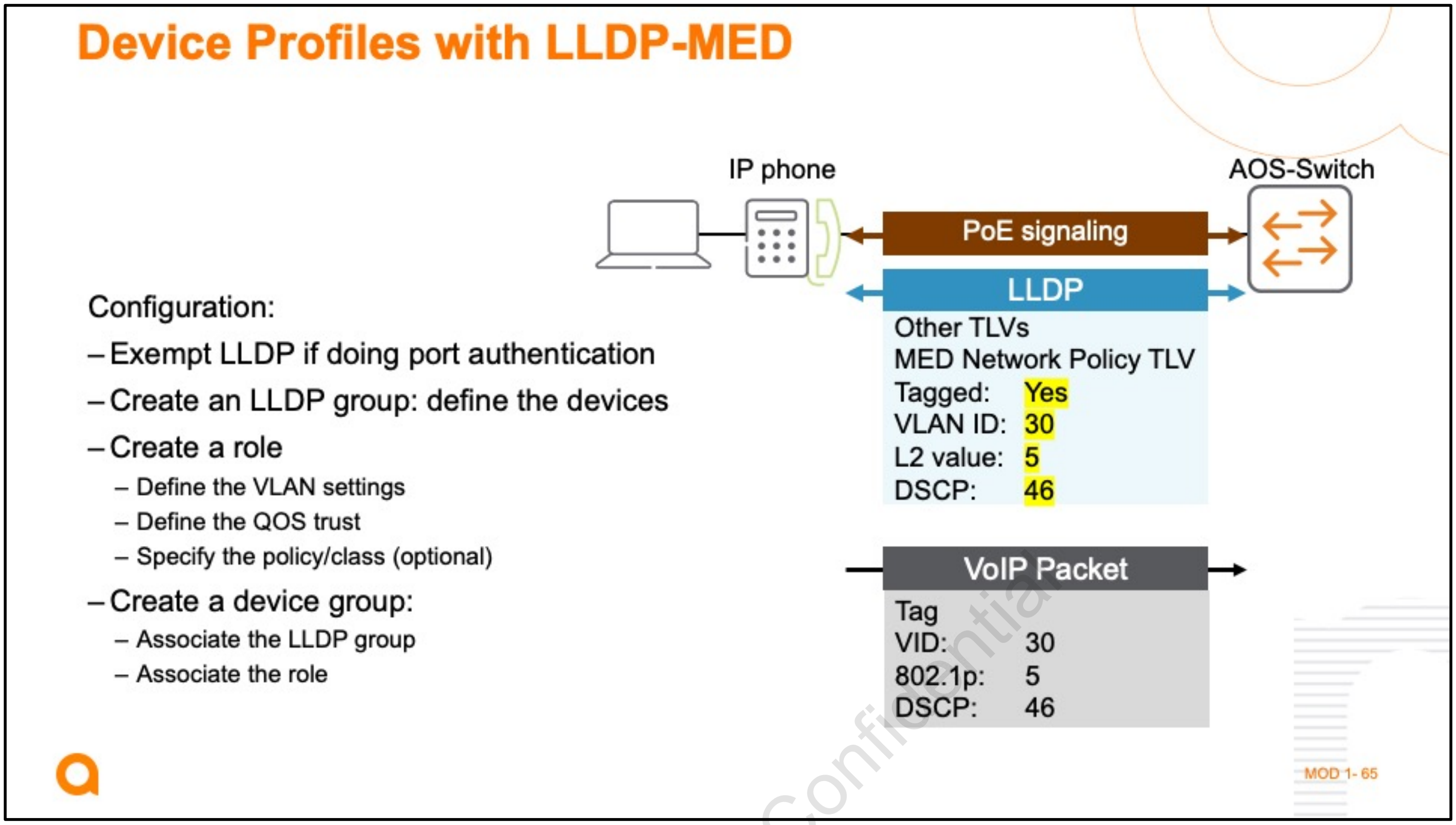
Policies: The policy associated with a role. This mode is set by the policy command.

PoE priority: The PoE priority applied to the interface. Supported values are high, low, and critical.

QoS Trust: QoS Trust mode applied to the interface. Supported values are dscp, tos, and none. This mode is set by the qos trust command.

• Jumbo frames: Enable or disable jumbo frame support for the interface. This mode is set by the mtu command.

The figure above lists the settings defined in a device profile, which is associated with the system description that contains “Aruba IAP”. Use the show lldp neighbor-info command of an associated device connected to the switch to determine what you should actually match on and include in your device profile matching parameters.



Device Profiles with LLDP-MED

Configuration:

- Exempt LLDP if doing port authentication
- Create an LLDP group: define the devices
- Create a role
  - Define the VLAN settings
  - Define the QOS trust
  - Specify the policy/class (optional)
- Create a device group:
  - Associate the LLDP group
  - Associate the role

Other TLVs
MED Network Policy TLV
Tagged: Yes
VLAN ID: 30
L2 value: 5
DSCP: 46

VoIP Packet
Tag
VID: 30
802.1p: 5
DSCP: 46

Suppose that an LLDP-MED capable phone boots up and connects to an AOS-CX LLDP-MED capable switch port. First the devices do a handshake to negotiate PoE, if required. The devices also exchange LLDP messages. The switch's LLDP message includes many TLVs, as you just learned.

If you are implementing authentication on the port, like 802.1X, MAC authentication, or captive portal, remember to allow LLDP:

```
switch(config)# interface <interface-ID>
switch(config-if)# aaa authentication port-access allow-lldp-bpdu
```

Creating an LLDP group

The LLDP group defines the devices to which you want to assign dynamic settings. To create an LLDP group, perform the following:

```
switch(config)# port-access lldp-group <LLDP-group-name>
switch(config-lldp-group)#  match <parameter> <parameter-value>
switch(config-lldp-group)#  ignore <parameter> <parameter-value>
```

The full match command syntax is as follows:

```
switch(config-lldp-group)# match sys-desc <description>
switch(config-lldp-group)# match sysname <system-name>
switch(config-lldp-group)# match vendo-oui <OUI-MAC-address>
switch(config-lldp-group)# match type <OUI-sub-type-key>
switch(config-lldp-group)# match value <OUI-sub-type-number>
```

This creates an LLDP group or modifies an existing LLDP group. An LLDP group is used to classify connected devices based on the LLDP type-length-values (TLVs) advertised by the device. A maximum of 32 LLDP groups can be configured on the switch. Each group accepts 64 match/ignore commands.

You can easily use the show lldp neighbor-info command to see what information a device is sending and thus what you should be matching. For example, to match on an Aruba IAP, the system description contains something like this: “ArubaOS (MODEL: 325), Version Aruba IAP”. So a corresponding match command to match on this would look like this:

```
switch(config-lldp-group)#  match sys-desc “Aruba IAP”
```

match and ignore commands are processed in sequence. Once a device matches on a match or ignore command, other commands in the device group are ignored. If you omit the sequence number, the command is added at the bottom of the LLDP group.

Note: CDP groups are also supported. See the Fundamentals Guide for configuring and using CDP groups.

Local User Roles (LURs)

The role defines the policies to use. To create a local role, perform the following:

```
switch(config)# port-access role <role-name>
switch(config-pa-role)# vlan access <VLAN-ID>
switch(config-pa-role)# vlan trunk native <VLAN-ID>
switch(config-pa-role)# vlan trunk allow <VLAN-ID-range>
switch(config-pa-role)# trust-mode <cos | dscp | none>
switch(config-pa-role)# associate policy <policy-name>
```

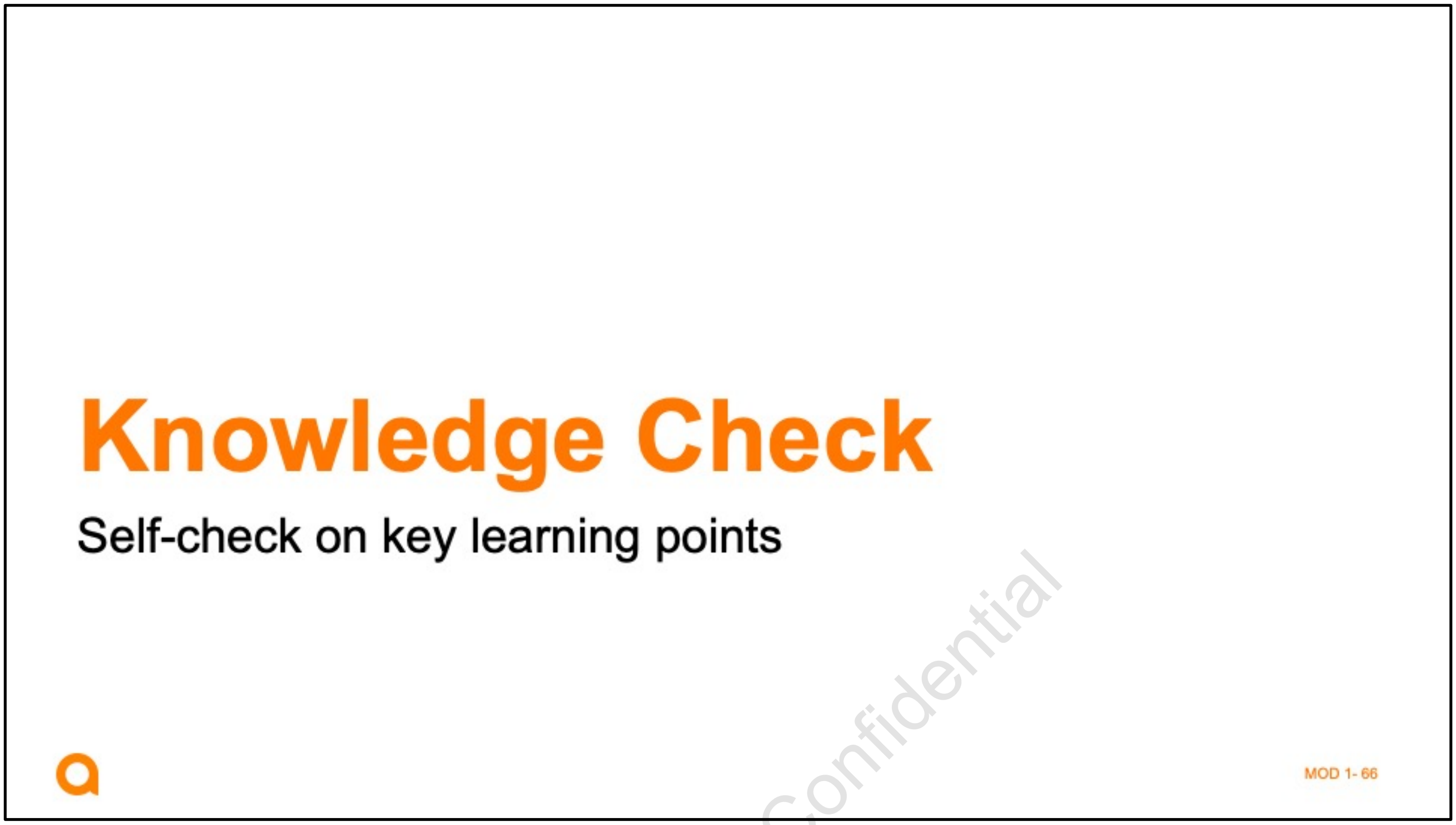
Creating a device profile

Next, you need to create a device profile and associate the LLDP group to it.

```
switch(config)# port-access device-profile <device-profile-name>
switch(config-device-profile)# associate lldp-group <name>
switch(config-device-profile)# associate role <name>
switch(config-device-profile)# enable
```

The role must be a local user role (downloadable user roles are not supported for this feature). Use the show port-access device-profile command to verify your configuration. The device profile is disabled by default, and must be enabled with the enable command. You can also disable it with the disable command.

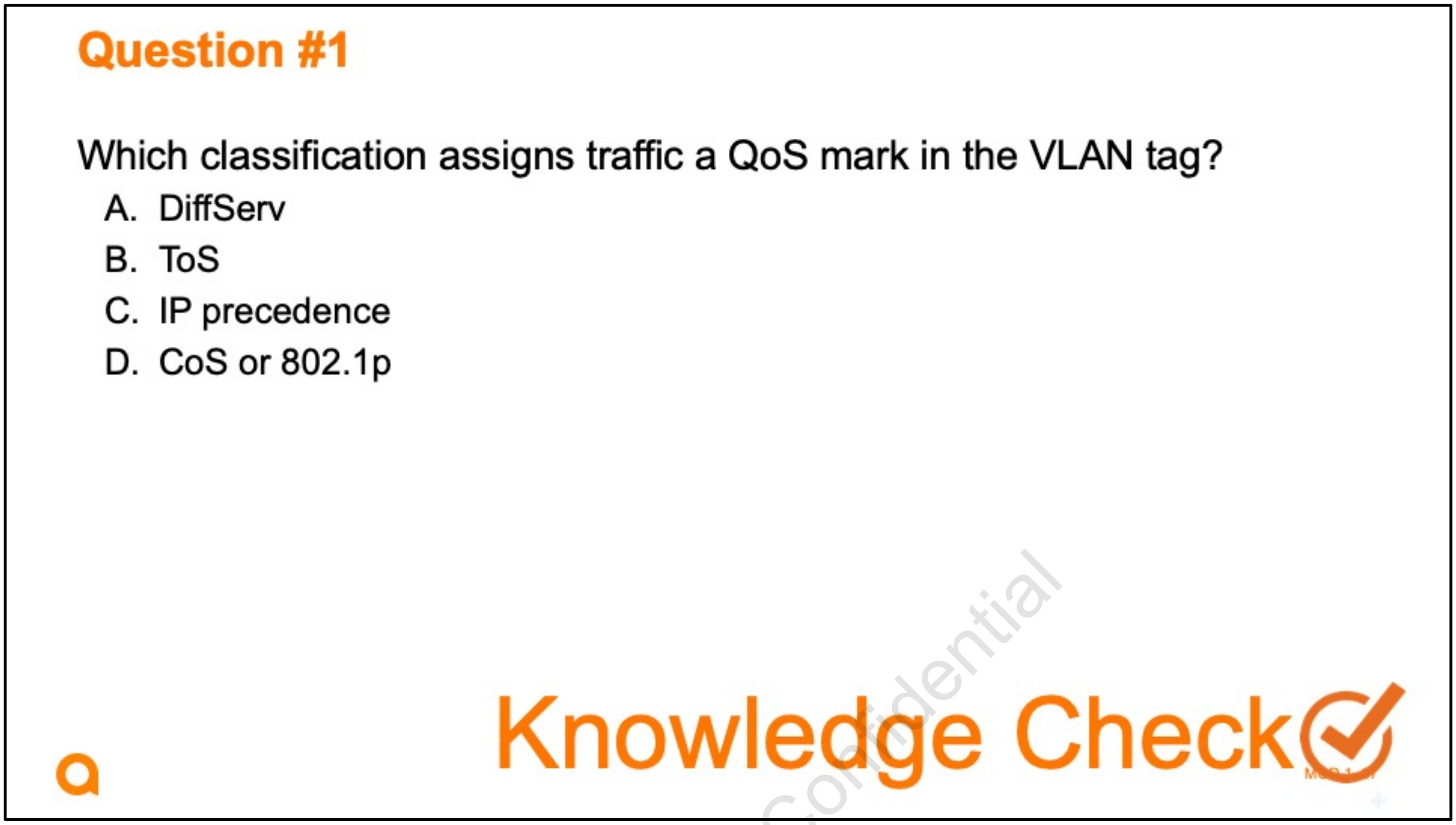
As you can see, you can specify VLAN and QoS polices. To define specific policy, you need to create class and policy configurations. A class specifies what to match on, and the policy map specifies the action(s) for the class map. The policy map is then referenced in the role with the associate policy command.



Knowledge Check

Self-check on key learning points

MOD 1- 66



Question #1

Which classification assigns traffic a QoS mark in the VLAN tag?

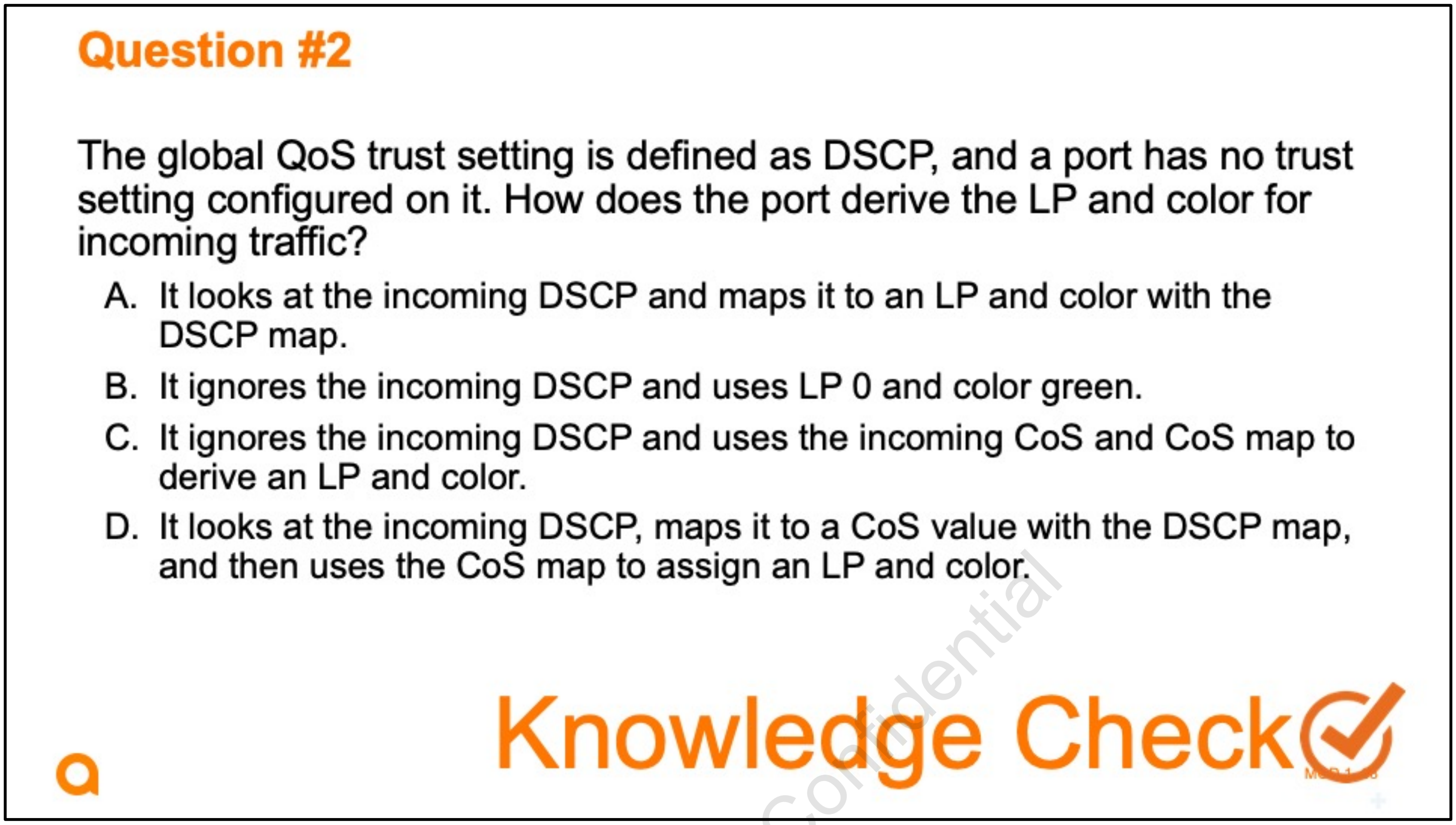
A. DiffServ

B. ToS

C. IP precedence

D. CoS or 802.1p

Knowledge Check



Question #2

The global QoS trust setting is defined as DSCP, and a port has no trust setting configured on it. How does the port derive the LP and color for incoming traffic?

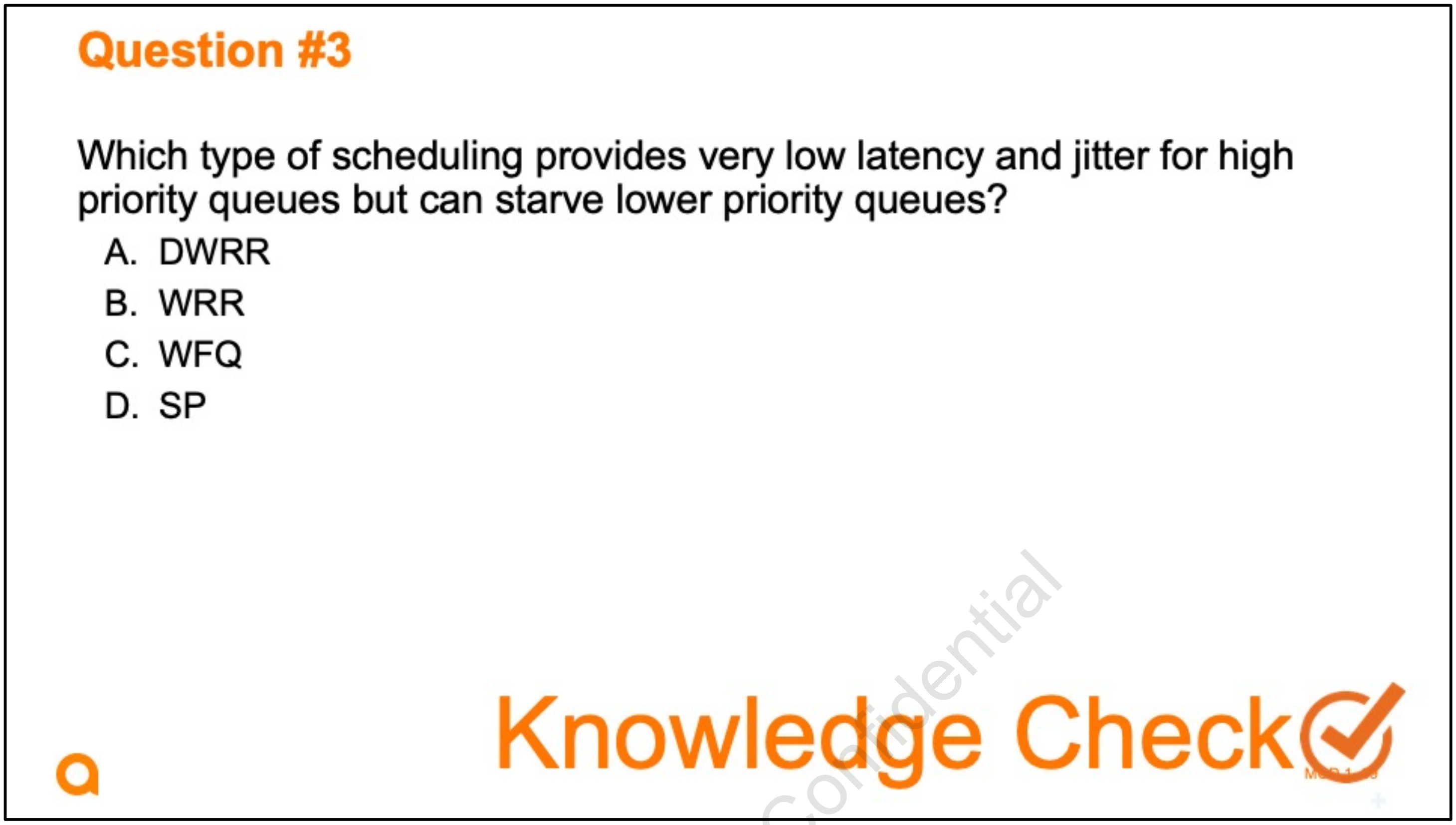
A. It looks at the incoming DSCP and maps it to an LP and color with the DSCP map.

B. It ignores the incoming DSCP and uses LP 0 and color green.

C. It ignores the incoming DSCP and uses the incoming CoS and CoS map to derive an LP and color.

D. It looks at the incoming DSCP, maps it to a CoS value with the DSCP map, and then uses the CoS map to assign an LP and color.

Knowledge Check



Question #3

Which type of scheduling provides very low latency and jitter for high priority queues but can starve lower priority queues?

A. DWRR

B. WRR

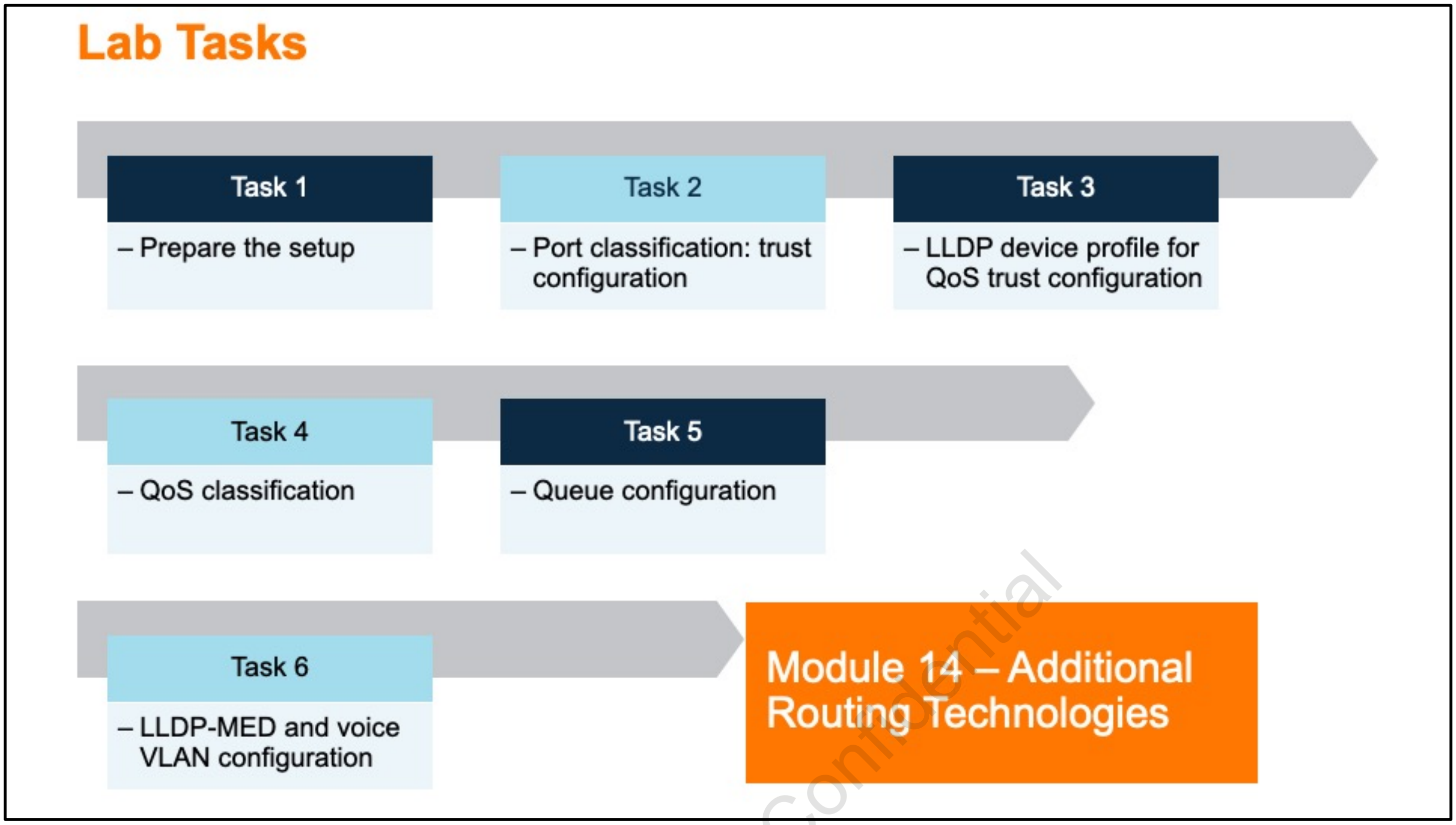
C. WFQ

D. SP

Knowledge Check

Lab Activity

Lab 6.3



Lab Tasks

Task 1

– Prepare the setup

Task 2

– Port classification: trust configuration

Task 3

– LLDP device profile for QoS trust configuration

Task 4

– QoS classification

Task 5

– Queue configuration

Task 6

– LLDP-MED and voice VLAN configuration

Module 14 – Additional Routing Technologies

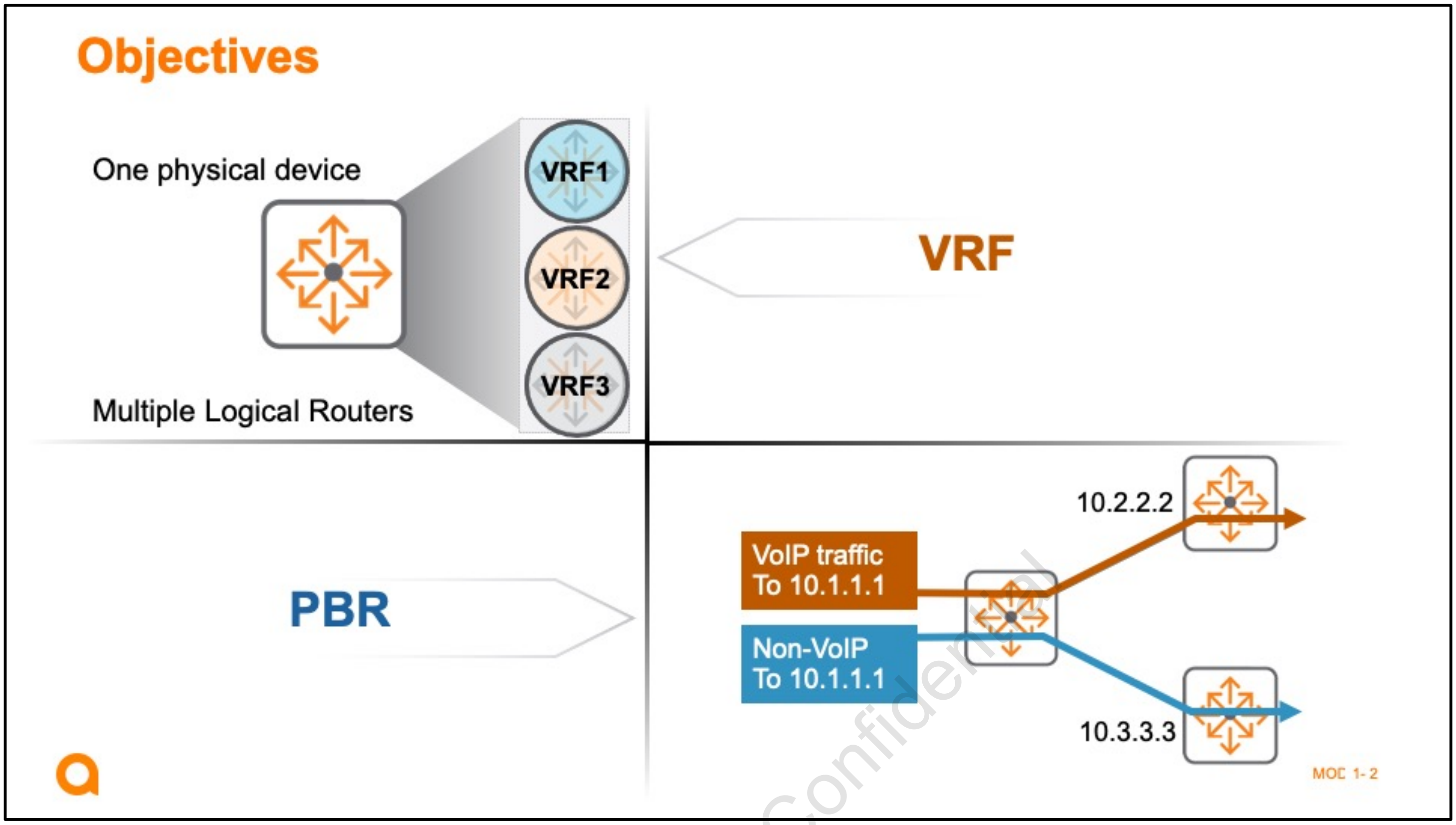
The figure provides a brief review of lab tasks. Please see your lab guide for details.

Rev # 20.21

Additional Routing Technologies

Implementing ArubaOS-CX Switching

Welcome back ladies and gentlemen. This is Module 14 – additional routing technologies.



Objectives

After completing this module, you should be able to:

Implement VRF to isolate routed traffic

Manipulate routing with policy-based routing (PBR)

Overview

VRF

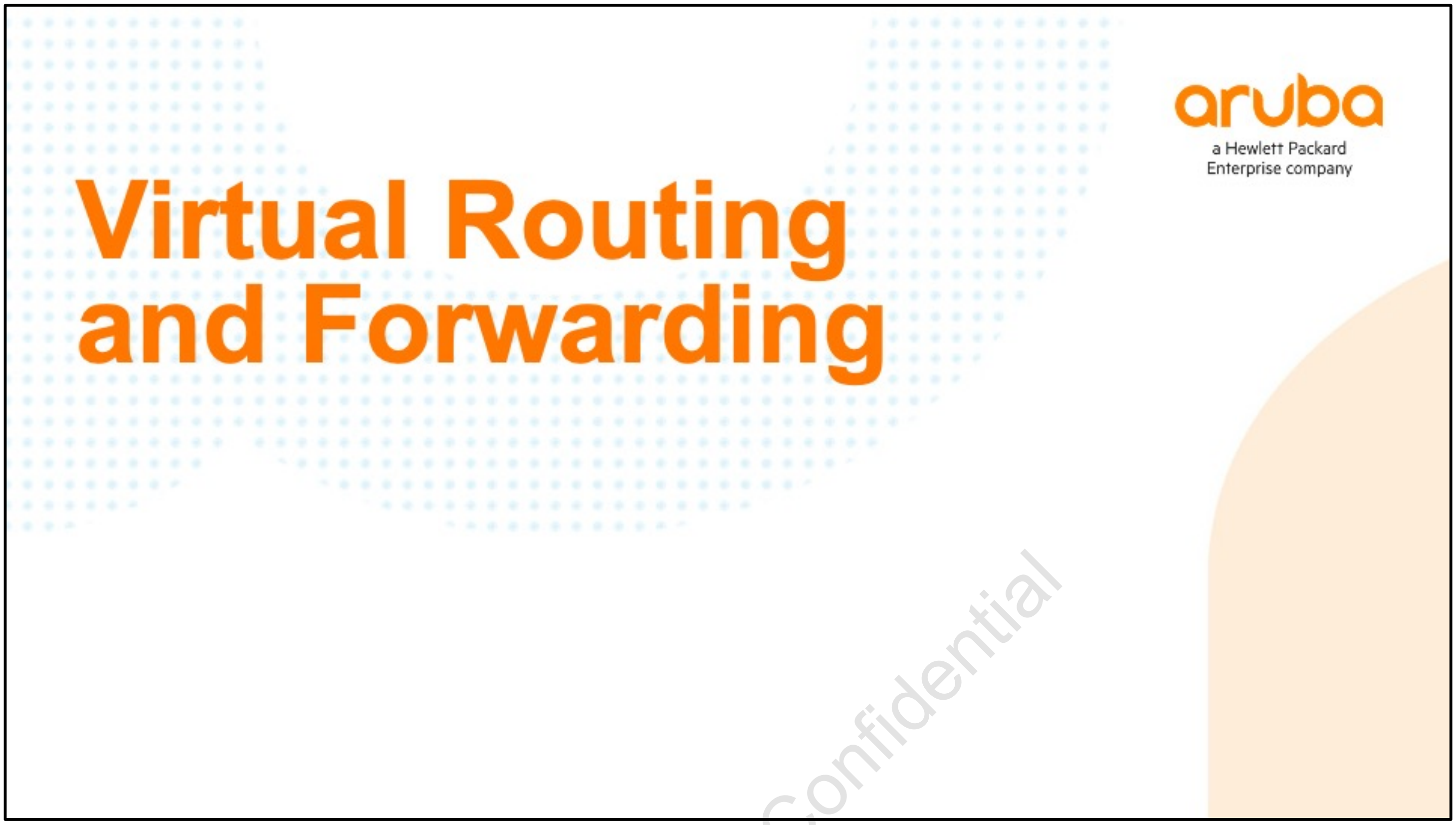
PBR

Lab Activity

MOD 1- 3

This module introduces two additional layer 3 routing topics:. First, you'll be introduced to virtual routing forwarding (VRF), which allows a company to logically separate their routing. This is a common practice by service providers, and allows duplicate addressing across the different VRFs. However, a company network sometimes needs this feature, commonly for security purposes.

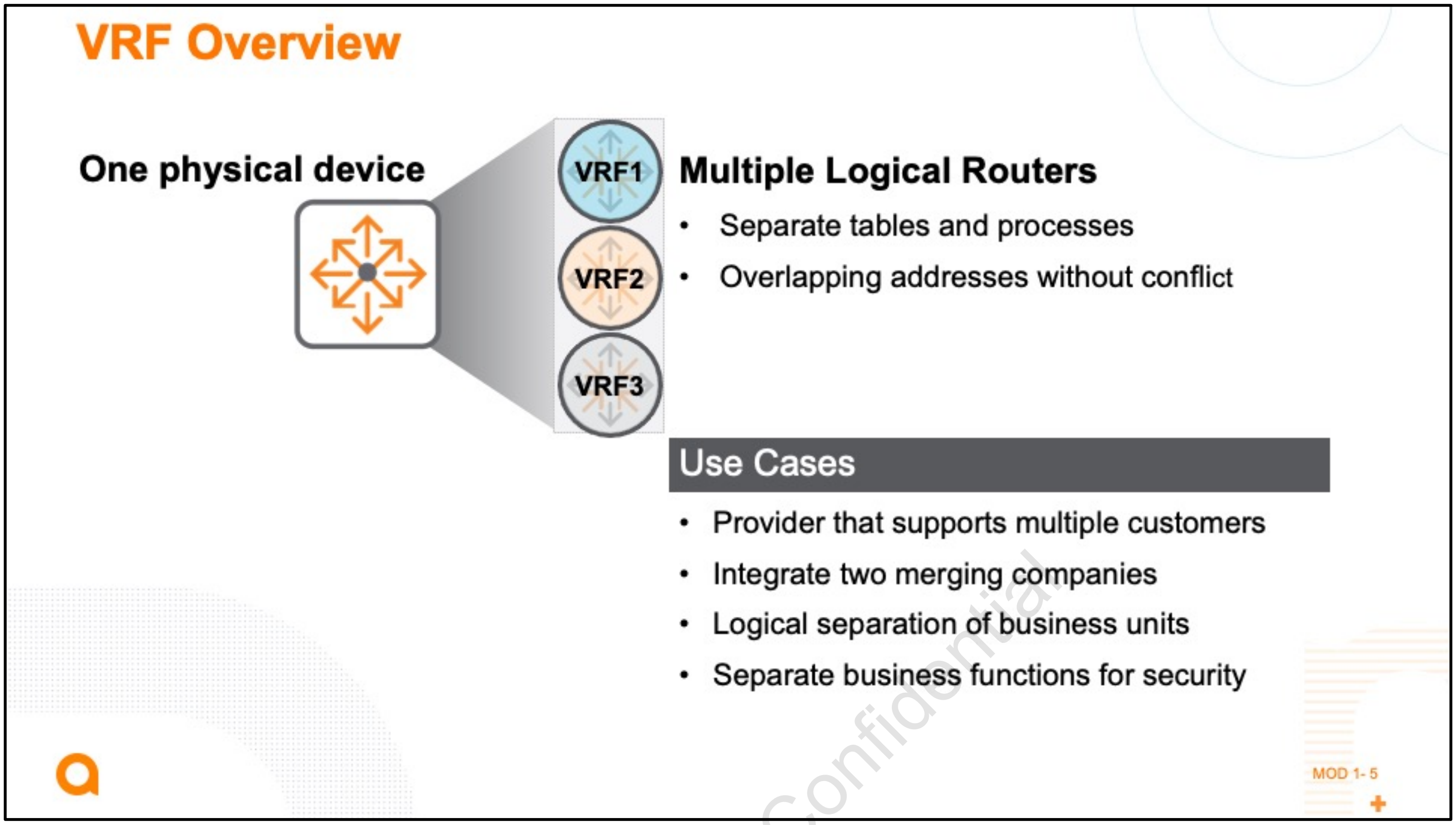
You will then be introduced to policy-based routing (PBR). PBR allows you to override the path in the routing table based on policy needs dictated by your company



Virtual Routing and Forwarding

aruba
a Hewlett Packard Enterprise company

We start with VRF.



VRF Overview

One physical device

Multiple Logical Routers

- Separate tables and processes
- Overlapping addresses without conflict

Use Cases

- Provider that supports multiple customers
- Integrate two merging companies
- Logical separation of business units
- Separate business functions for security

MOD 1- 5

Virtual Routing and Forwarding (VRF) allows a single physical device to host multiple logical routers, with complete L3 isolation. Because the device maintains separate route processes and tables for each VRF, they can have overlapping address space. Network functionality is improved because network paths can be segmented without requiring multiple routers.

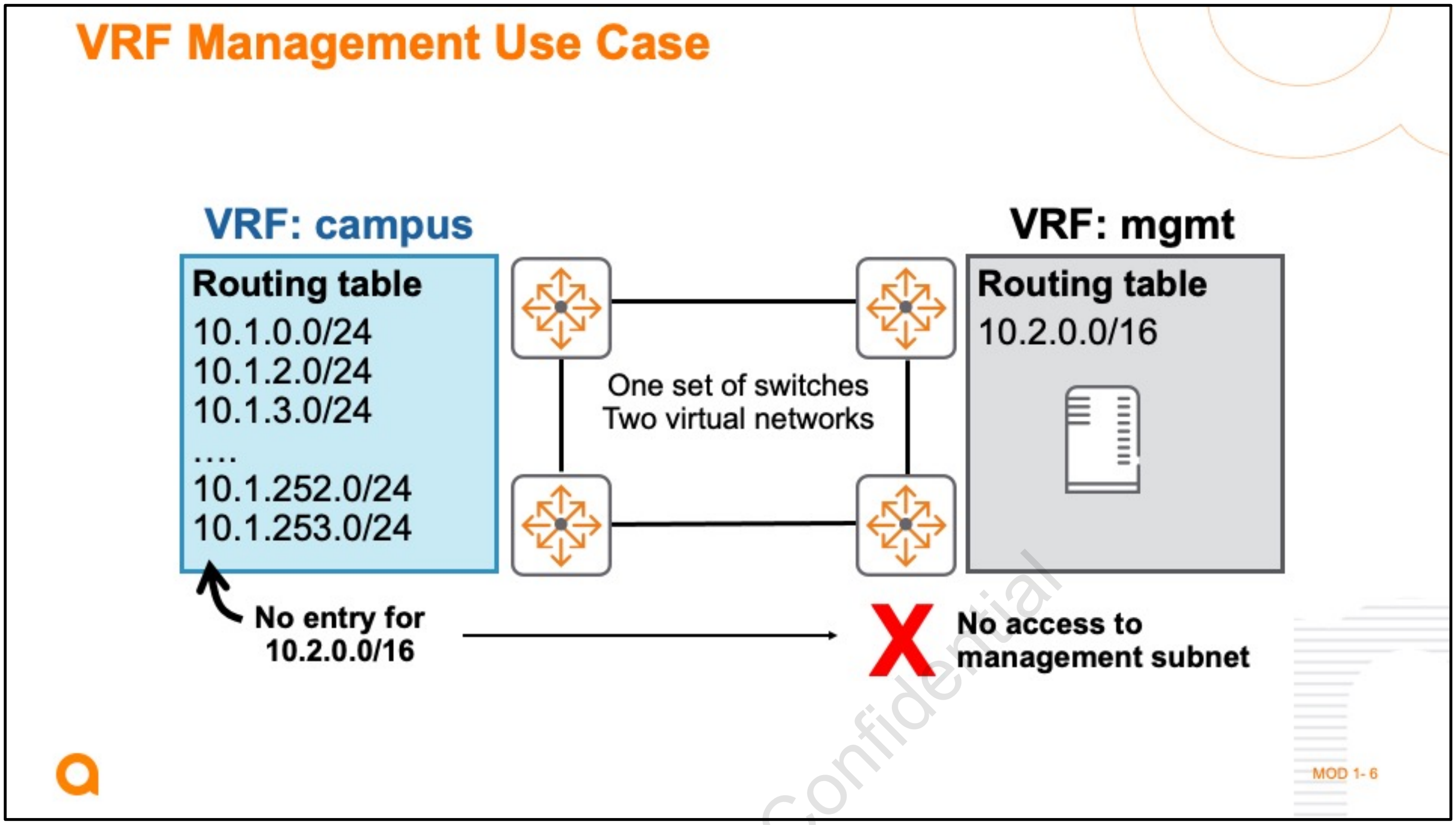
| Perhaps the scenario shown in the figure represents a single service provider, hosting separate environments for three customers. Customers can use any address space they choose, without fear of conflicting with other customers.

Another use case is when you are merging with another organization. You can physically connect the two organizations without having to worry about addressing conflicts. As you normalize the address space between the two entities, you can eliminate the VRF configuration as appropriate.

Another common use case is to maintain logical separation of business units. For example, a banking company might maintain separation between their retail, lending/mortgages, investments, and internal applications.

A common use case is to separate business functions. Place network management functions in one VRF, with the rest of the network in another VRF. Management network routes are unknown to the rest of the network, providing a logical separation between the two – you get enhanced security. Or in a data center place public resources in one VRF, with private or internal services in a separate VRF. Even though the both sets of services are in the same data center, they are logically separated since their routes are in different routing tables, thus enhancing security.

Let’s analyze some scenarios in more detail.



VRF Management Use Case

VRF: campus

Routing table
10.1.0.0/24
10.1.2.0/24
10.1.3.0/24
….
10.1.252.0/24
10.1.253.0/24

One set of switches
Two virtual networks

VRF: mgmt

Routing table
10.2.0.0/16

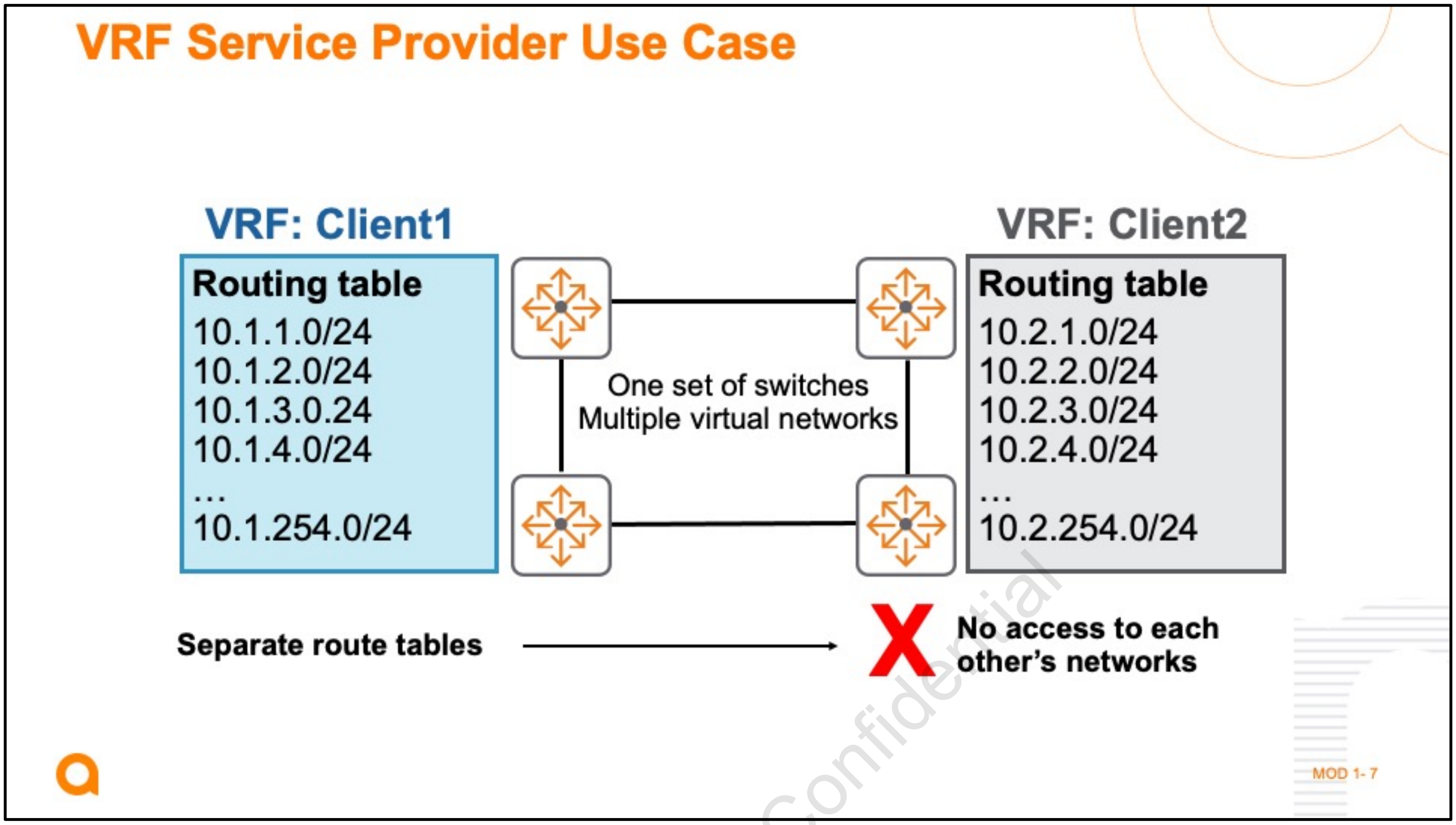
No entry for
10.2.0.0/16

No access to
management subnet

MOD 1- 6

In this example, network management traffic is logically separated from the rest of the campus traffic. One set of switches support two separate VRF instances, each with their own, separate set of routing tables. All campus users and services connect to interfaces that have been placed in the Campus VRF. This logically separates the campus management and data plane functions. Since nobody in the campus VR can “see” the management VRF routes, they can’t reach them. You get solid security between the two domains – no access-lists required.

Of course, the management VRF does not have campus routes in its routing table and so devices in the management VRF cannot reach resources in the campus VRF.



VRF Service Provider Use Case

VRF: Client1

Routing table
10.1.1.0/24
10.1.2.0/24
10.1.3.0.24
10.1.4.0/24
…
10.1.254.0/24

VRF: Client2

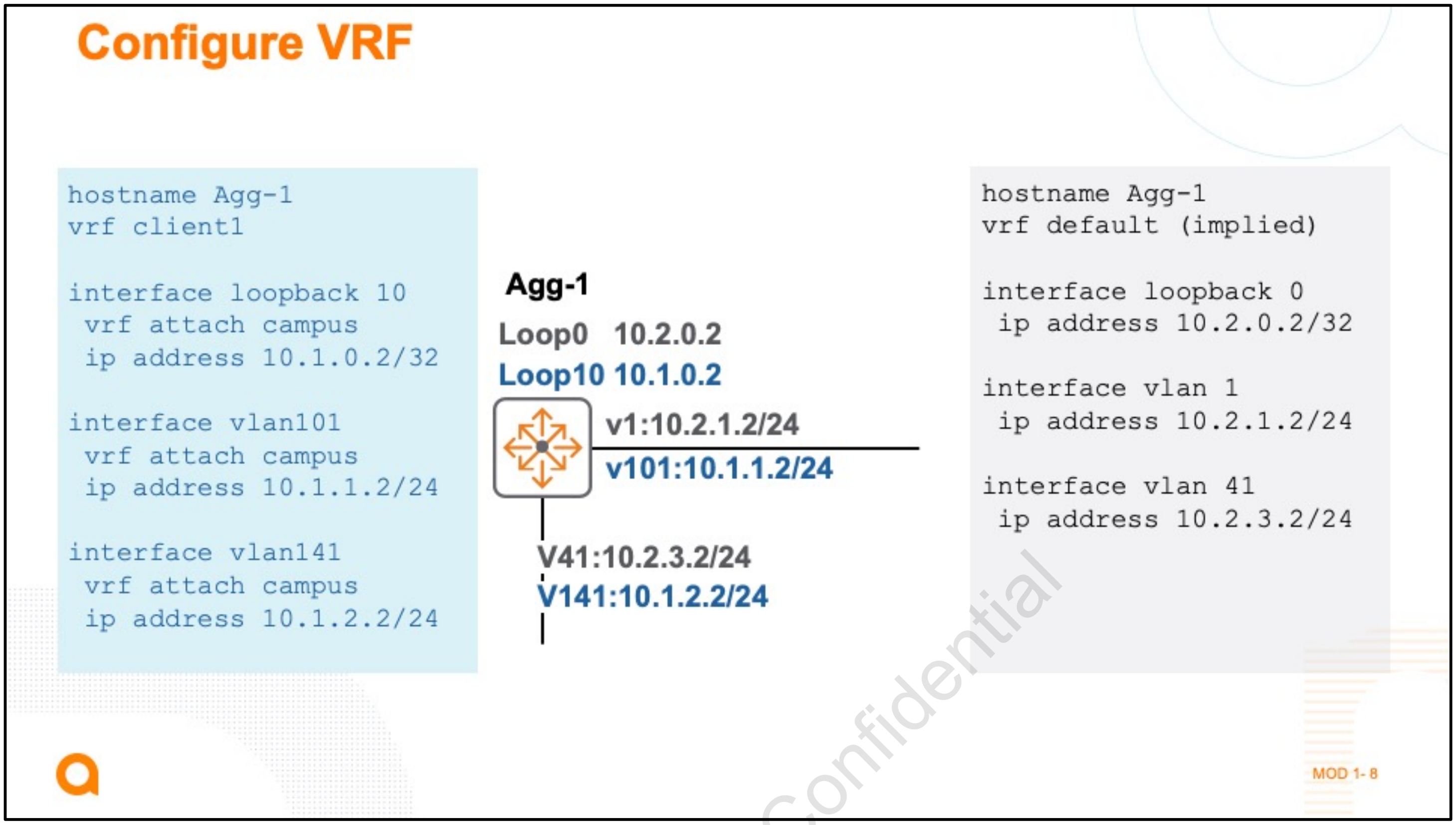
Routing table
10.2.1.0/24
10.2.2.0/24
10.2.3.0/24
10.2.4.0/24
…
10.2.254.0/24

MOD 1- 7

In this example, a service provider uses AOS-CX switches and VRF to logically separate two companies, Client1 and Client2. Notice that even though both companies have an overlapping route, the VRF implementation logically separates the two companies so that they cannot directly interact with each other.

By default, the two companies require an intermediate firewall to interact with each other, using address translation to deal with the overlapping address space. The two companies would have their own public address space and thus would translate their internal address space to a unique public address space in order to interact with other resources in the data center or on the Internet.

Note: Different interfaces must be used for the two VRFs (a VRF cannot share an interface.



Configure VRF

```
hostname Agg-1
vrf client1

interface loopback 10
 vrf attach campus
 ip address 10.1.0.2/32

interface vlan101
 vrf attach campus
 ip address 10.1.1.2/24

interface vlan141
 vrf attach campus
 ip address 10.1.2.2/24
```

```
hostname Agg-1
vrf default (implied)

interface loopback 0
 ip address 10.2.0.2/32

interface vlan 1
 ip address 10.2.1.2/24

interface vlan 41
 ip address 10.2.3.2/24
```

To create a VRF, use the global config command vrf <vrf-name>. The VRF name can be up to 32 alphanumeric characters. You are allowed a total of 65 VRFs on your routing switch. By default, two exist on the switch:

default: used if you do not specify a VRF with a configuration command

mgmt: used for out-of-band management (OBBM)

This means that you can create an additional 63 VRFs. And remember, an interface can only belong to one VRF.

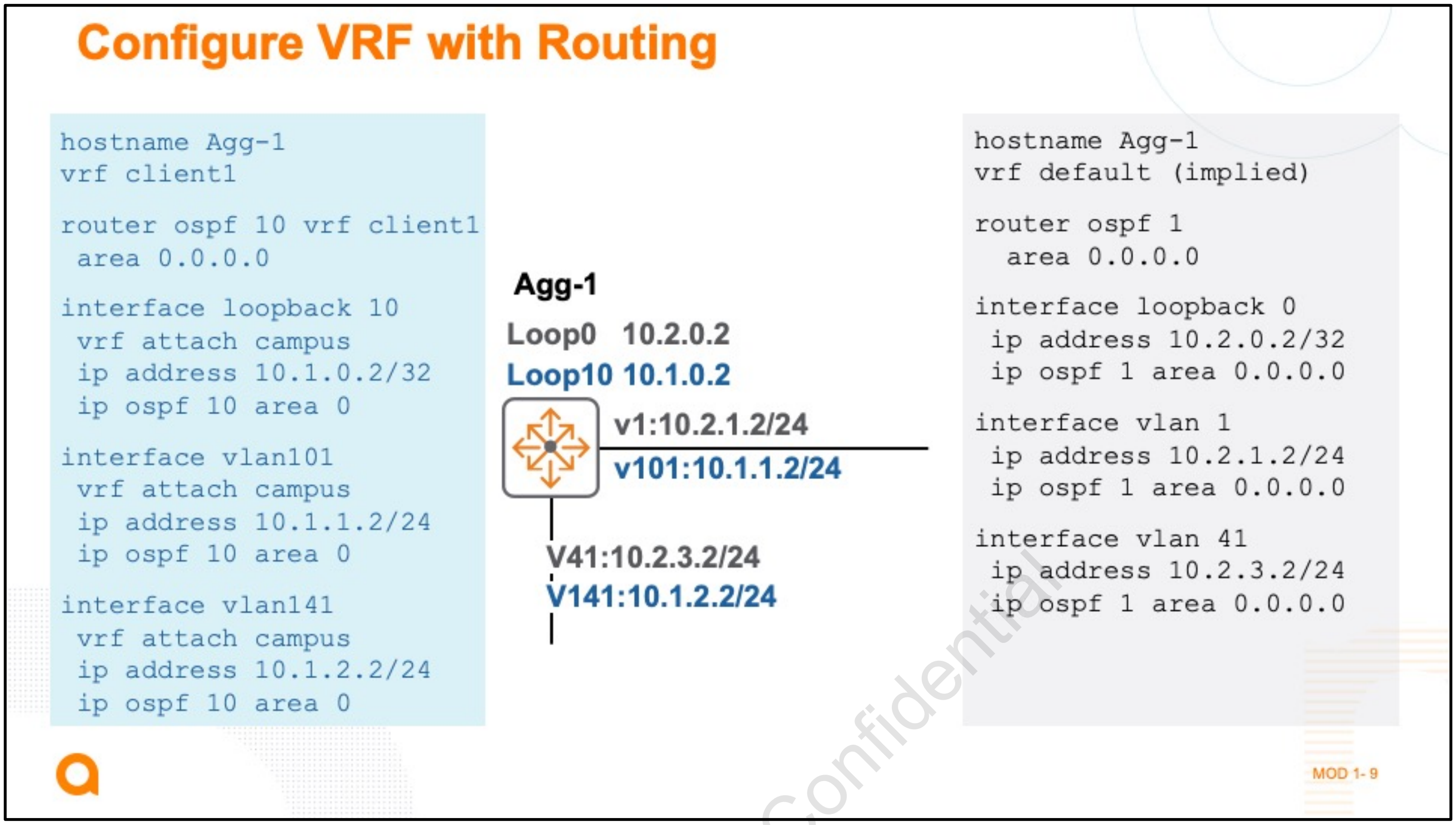
The figure shows an example of one physical infrastructure that supports two separate VRFs. The blue configuration to the left is for VRF “client1”. This configuration is reflected in the network diagram – just focus on the blue addresses and ignore the grey addresses.

Notice that each interface is configured with vrf attach campus. This is how you tell the router which interfaces belong to which VRF. If you enter the command no vrf attach campus, the interface is returned to the default VRF, and all interface configuration is removed. This is true anytime you move an interface to a different VRF.

The configuration to the right is for VRF default, also reflected in the network diagram. Just focus on the grey addresses and ignore the blue addresses. The command vrf default is shown but need not be configured – it is the default.

Note: This configuration matches the Service Provider use case you just reviewed. You would probably use VRF “client1” and VRF “client2”. The default VRF was used for client2 to ensure you understand that all interfaces are in the default VRF unless otherwise specified.

You have now split one physical device into two logical routers, but no routing protocol is configured. Let’s add OSPF routing to this deployment.



Configure VRF with Routing

```
hostname Agg-1
vrf client1

router ospf 10 vrf client1
 area 0.0.0.0

interface loopback 10
 vrf attach campus
 ip address 10.1.0.2/32
 ip ospf 10 area 0

interface vlan101
 vrf attach campus
 ip address 10.1.1.2/24
 ip ospf 10 area 0

interface vlan141
 vrf attach campus
 ip address 10.1.2.2/24
 ip ospf 10 area 0
```

```
hostname Agg-1
vrf default (implied)

router ospf 1
  area 0.0.0.0

interface loopback 0
 ip address 10.2.0.2/32
 ip ospf 1 area 0.0.0.0

interface vlan 1
 ip address 10.2.1.2/24
 ip ospf 1 area 0.0.0.0

interface vlan 41
 ip address 10.2.3.2/24
 ip ospf 1 area 0.0.0.0
```

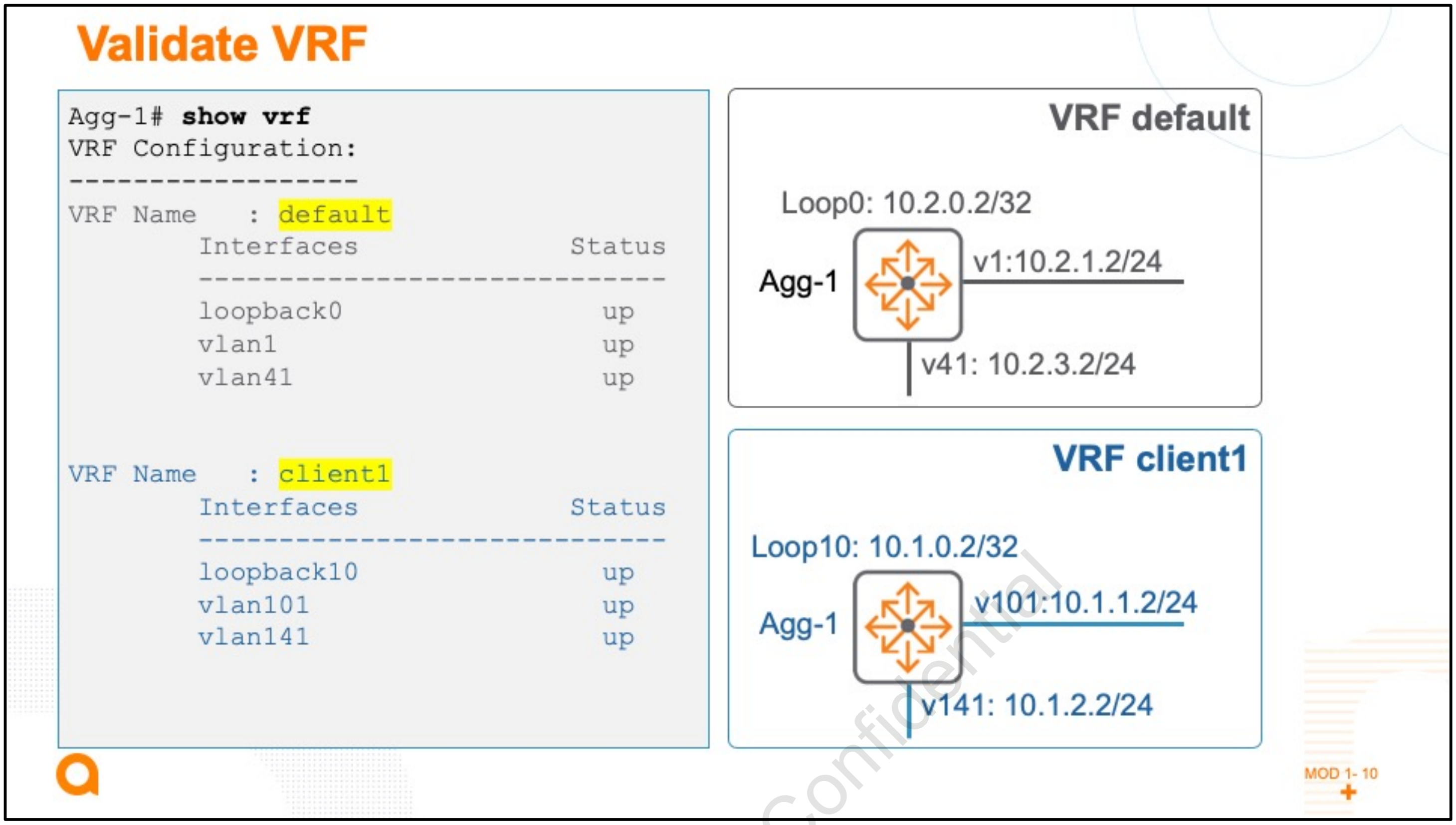
MOD 1- 9

Many routing switch commands allow you to specify the VRF name to which the command applies. The figure shows the previous example, but with OSPF configuration added to the blue client1 configuration on the right. The syntax to configure OSPF is as typical, except for the addition of the VRF name at the end.

Many configuration and show commands allow you to reference a VRF name and thus compartmentalize your routing functions. If you do not use an optional VRF name, then the command applies to the default VRF.

Consider a static routing example, with the command ip route 10.254.0.0/16 10.5.5.5. This tells the router to use next-hop router 10.5.5.5 to reach destination network 10.254.0.0/16. Since you did not specify a VRF, this route is placed in the default VRF route table.

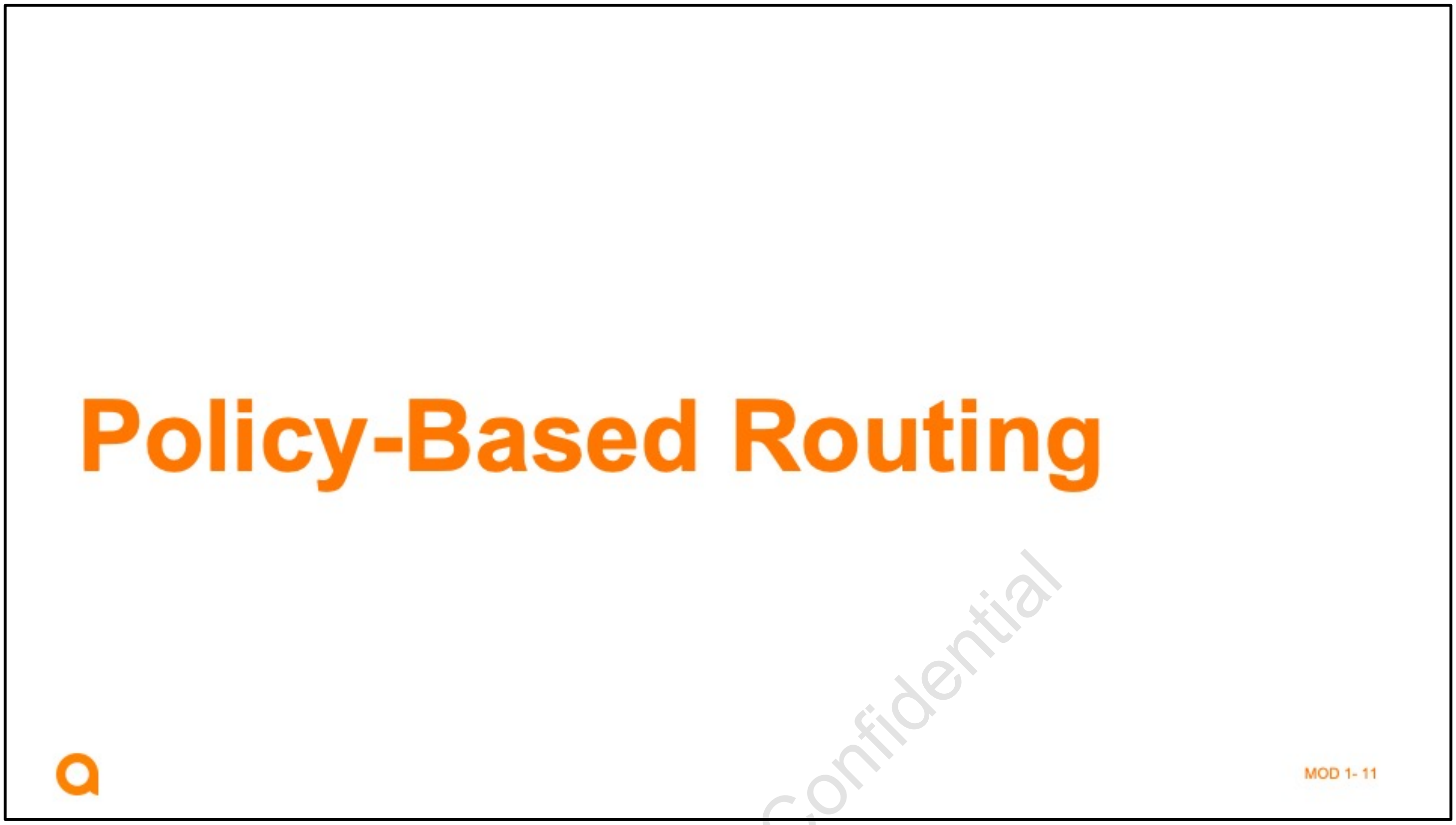
To place this route in the client1 VRF, use ip route 10.254.0.0/16 10.5.5.5 vrf client1



Validate VRF

```
Agg-1# show vrf
VRF Configuration:
------------------
VRF Name    : default
        Interfaces                 Status
        -----------------------------
        loopback0                    up
        vlan1                        up
        vlan41                       up


VRF Name    : client1
        Interfaces                 Status
        -----------------------------
        loopback10                   up
        vlan101                      up
        vlan141                      up
```

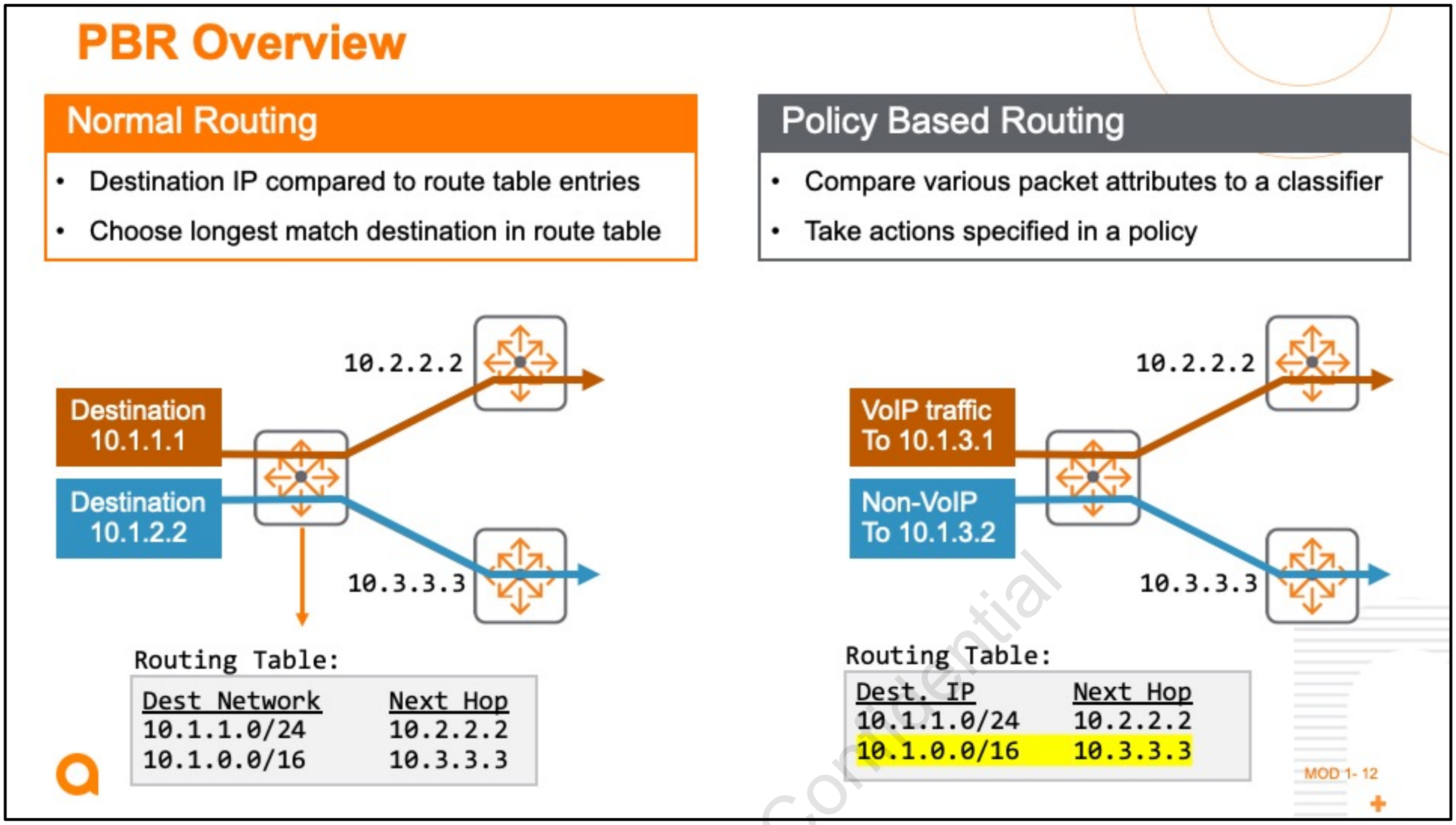
The figure shows how to validate which VRFs exist, and which interfaces are associated with each VRF. To look at each VRF separately, you could use show vrf default and show vrf client1.



Policy-Based Routing

MOD 1- 11

Now you explore Policy-Based Routing (PBR) - to override normal destination-based routing entries learned by static, OSPF, or BGP routes.



PBR Overview

Normal Routing

- Destination IP compared to route table entries
- Choose longest match destination in route table

Policy Based Routing

- Compare various packet attributes to a classifier
- Take actions specified in a policy

Routing Table:

Dest Network	Next Hop
10.1.1.0/24	10.2.2.2
10.1.0.0/16	10.3.3.3

Routing Table:

Dest. IP	Next Hop
10.1.1.0/24	10.2.2.2
10.1.0.0/16	10.3.3.3

MOD 1- 12

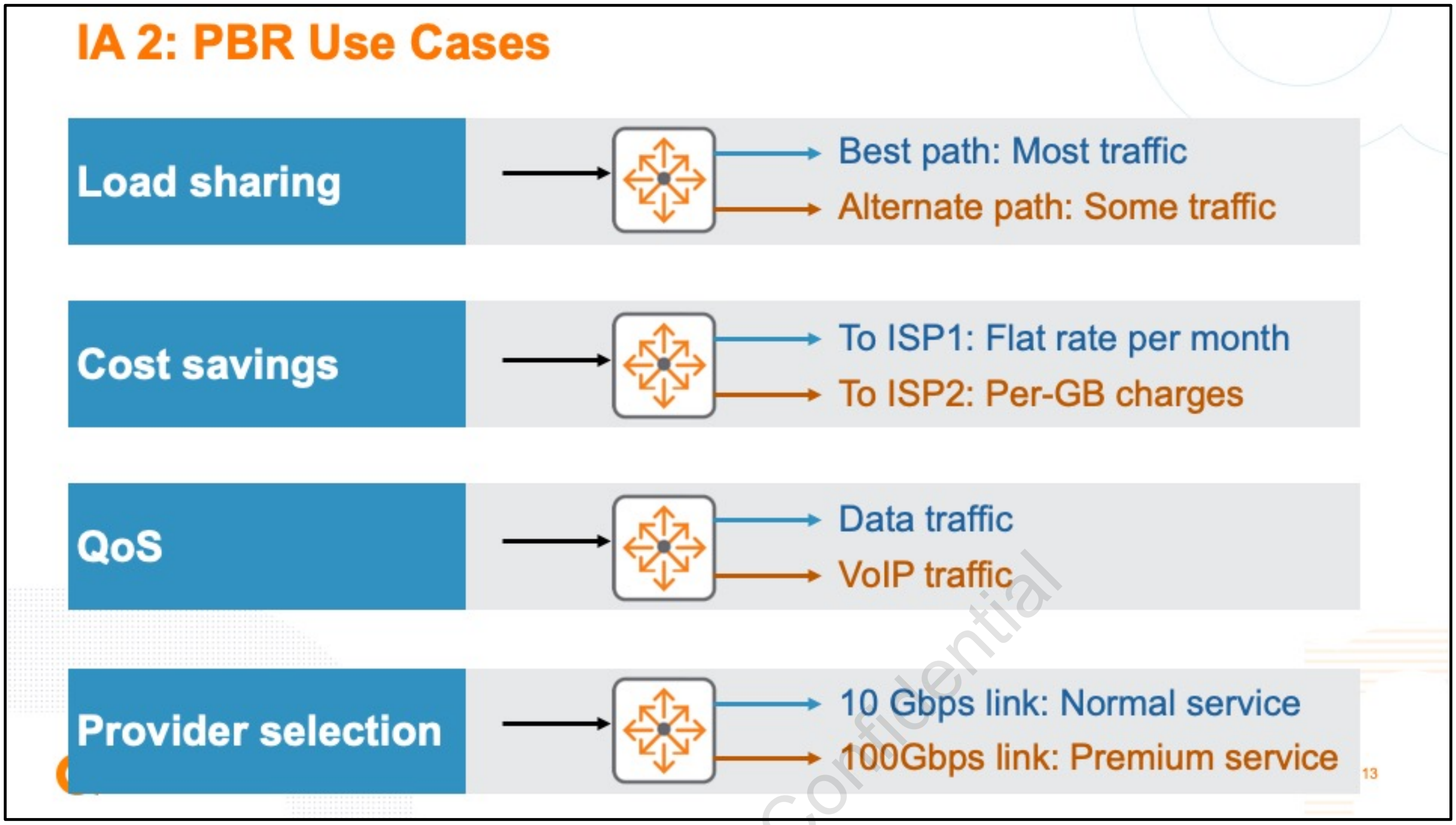
The figure compares normal routing to Policy Based Routing (PBR).

Normal routing is purely destination-based. As each router receives an inbound packet, it compares the packet's destination IP address to the route table, then chooses the entry with the longest match as the best path. In the example, the packet destined to 10.1.1.1 matches both entries in the route table, but only 16 bits match in the top entry, while 24 bits match in the second entry. So, the router routes this packet via next hop 10.2.2.2. Next, a packet arrives destined to 10.1.2.2. This destination only matches the top entry, and so 10.3.3.3 is the next hop.

| PBR lets you manipulate the path of a packet based on various packet attributes. Packets must be destined to a valid subnet on your network. As each router receives an inbound packet, it compares appropriate header and tag information to a classifier that you configure. The router then takes action on this matching traffic based on a policy. Matching traffic with the same destination can be routed over different paths, to improve traffic handling, performance, and/or security for various traffic types.

In the example, both packets are for the same destination subnet. However, you have created a classifier that matches on the range of UDP ports used for Voice over IP (VoIP) traffic – perhaps UDP ports 5004-5065. This traffic is routed over its own path via destination 10.2.2.2. Note that this action overrides what normal routing would do by default – the ingress router only has next-hop 10.3.3.3 for 10.1.0.0/16. However, 10.2.2.2 is a valid path, and it is directly connected to the ingress router, as required by PBR. In other words, recursive lookups are not supported.

All other, non-VoIP traffic does not match the classifier, and so is routed as normal, via 10.3.3.3. PBR only affects the local router's routing decision. You might have to implement the PBR policy on multiple routing devices in order to achieve the desired outcome if multiple hops exist between a source and destination.



IA 2: PBR Use Cases

By default, a routing device uses its routing table to determine which path to take when reaching a destination. There are times, though, where a company needs to override the default routing behavior based on traffic types or network conditions. The figure summarizes use cases – click on “+” icon each to explore.

Load sharing

Imagine that you have two links for redundancy for an OSPF network, but one link has a better cost. The higher cost link is available, but is not used, unless the better cost link fails. Configure PBR to use the secondary link for certain traffic and thus implement a limited form of load balancing.

Cost savings

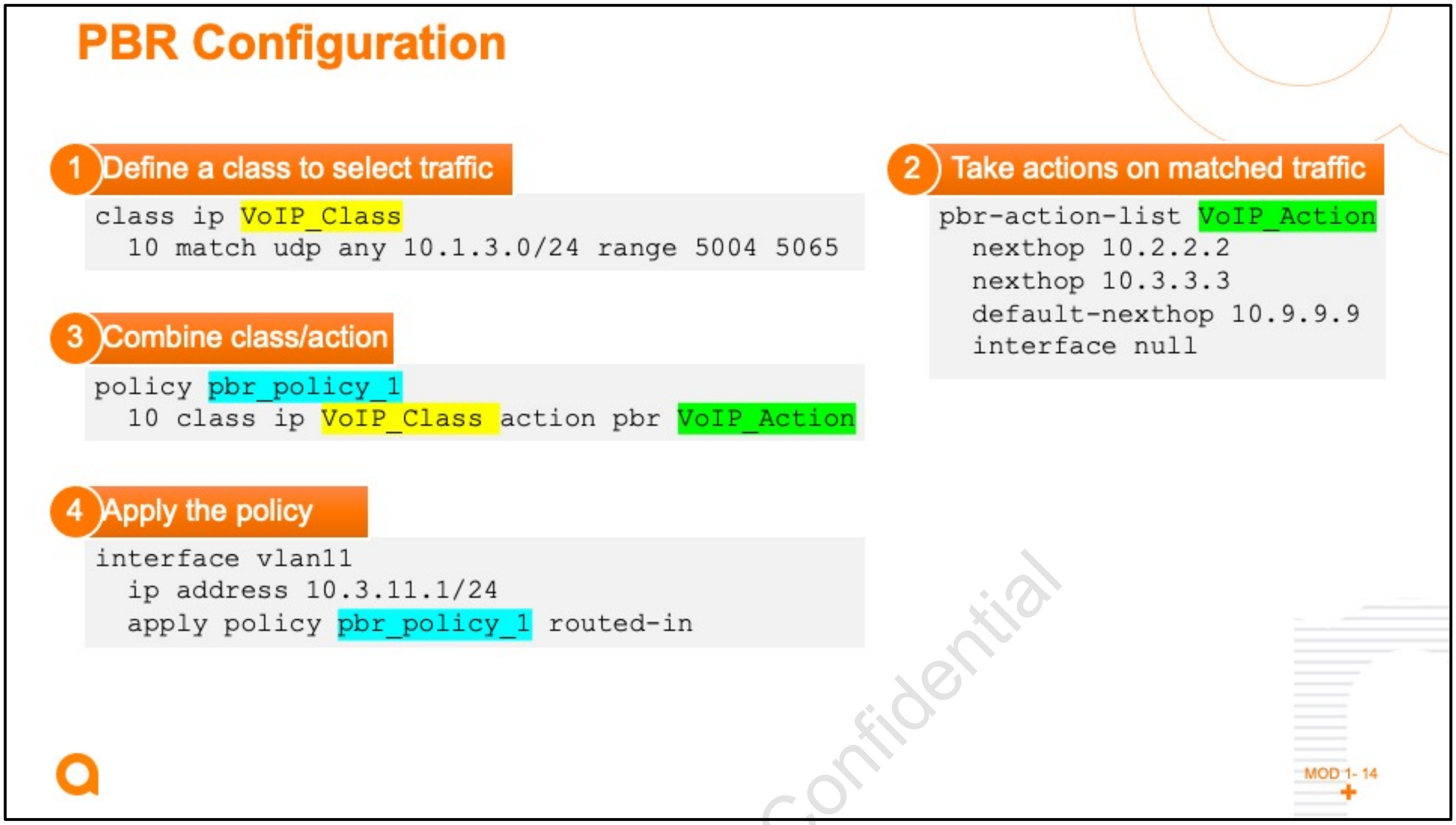
Suppose that you have two Internet connections via two different providers: ISP1 charges a flat rate and ISP2 charges a flat rate with a usage cap. Once you reach the cap, you are charged per gigabyte sent or received. Use PBR to ensure that even if the ISP2 has better metric routes, that ISP is not used unless ISP1 is unreachable.

QoS

Perhaps you have two equal-cost paths for some destination. You want to send delay and jitter-sensitive VoIP traffic on one link, and all other traffic over the other link. You maintain this redundancy while logically separating the traffic. VoIP traffic gets a link all to itself, but you must still configure QoS in case the VoIP link fails.

Provider selection

You might need to control traffic paths based on the source IP address, as opposed to normal destination-based routing. Imagine that you work for a service provider offers normal and premium-level services. The backbone has two sets of connections: 10 Gbps and 100 Gbps. Use PBR to ensure that traffic sourced from premium-tier customers traverse the 100 Gbps connections, while normal customer traffic uses 10 Gbps connections.



PBR Configuration

1 Define a class to select traffic

```
class ip VoIP_Class
  10 match udp any 10.1.3.0/24 range 5004 5065
```

2 Take actions on matched traffic

```
pbr-action-list VoIP_Action
  nexthop 10.2.2.2
  nexthop 10.3.3.3
  default-nexthop 10.9.9.9
  interface null
```

3 Combine class/action

```
policy pbr_policy_1
  10 class ip VoIP_Class action pbr VoIP_Action
```

4 Apply the policy

```
interface vlan11
  ip address 10.3.11.1/24
  apply policy pbr_policy_1 routed-in
```

MOD 1- 14

The example in the figure provides an overview of PBR configuration.

1. Classifier

First you define a classifier – to select which traffic shall have some special actions applied to it. In the example, you create a class called “VoIP_Class”. This class shall include UDP traffic from any source to destination 10.1.3.0/24, with a destination UDP port between 5004 and 5065 inclusive.

2. Action list

Next, define a PBR action list – to control how some traffic class is to be treated. In the example, you created a list named “VoIP_Action”, which contains four possible actions in an ordered list. You can have a maximum of eight action entries in an action list, 512 total entries across the entire device.

Action List entries are processed top-down, like an ACL. The first entry indicates that the next hop for matched traffic is 10.2.2.2. If that next hop is not available, perhaps due to a link, interface, or router down condition, then the next hop is 10.3.3.3. If that next hop router is also down, then use a default next hop of 10.9.9.9. And finally, if that next hop is down, silently discard the packet. In other words, forward it out the “null” interface – the proverbial bit bucket.

The example shows three of the four possible actions. The other one is interface tunnel, which sends traffic over some GRE, 6in4, or 6in6 tunnel. The example does not show the default sequence numbers, which are automatically assigned based on order of entry – in increments of 10. You can configure your own sequence numbers on action list entries if desired.

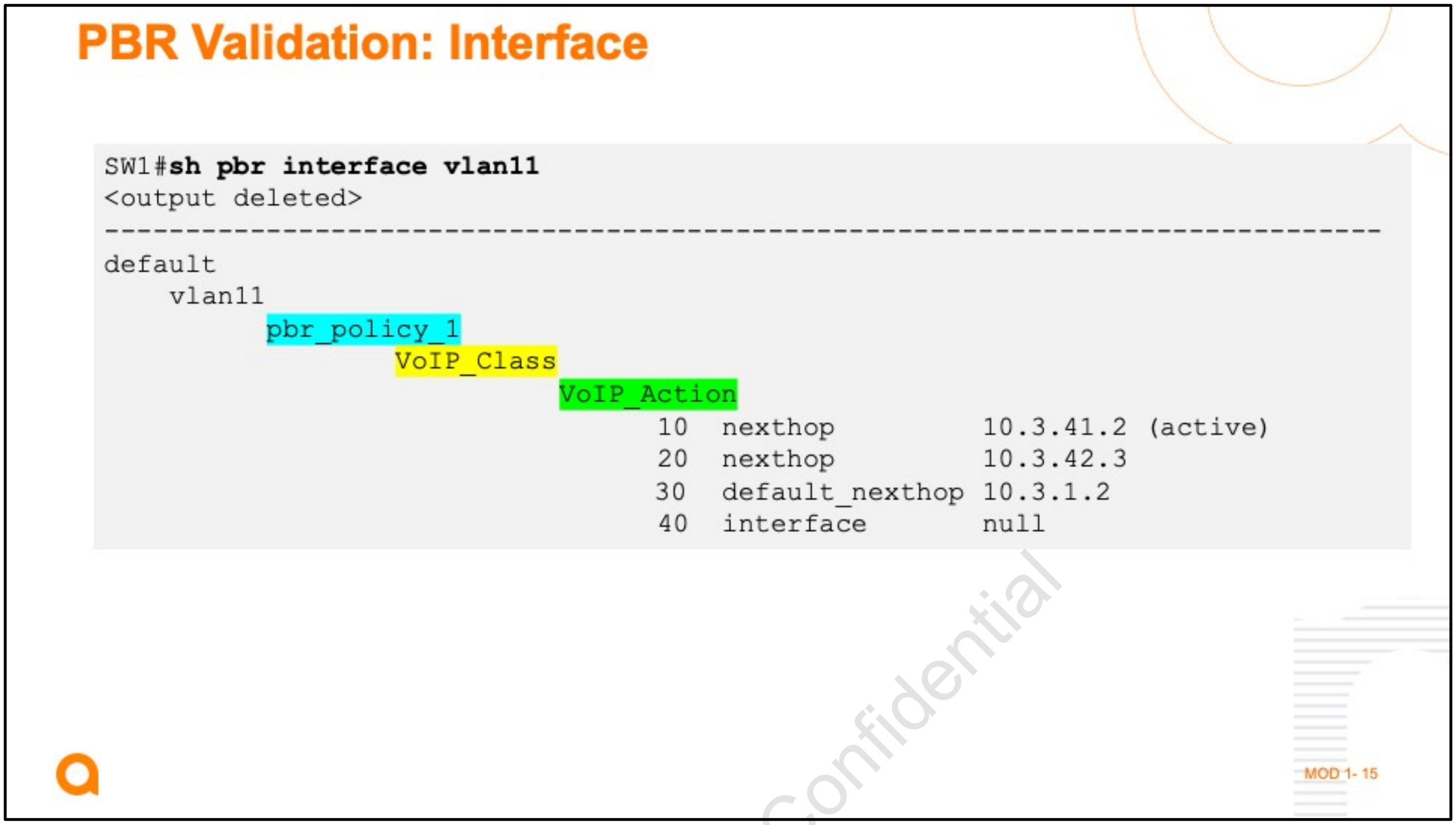
Understand that next hop actions must point to directly connected IP addresses – PBR does not support recursive lookups to reach non-directly connected next hop addresses.

3. Create a policy to combine classes and actions

Now you create a policy to associate each class with an appropriate action list. The example associates the VoIP_Class with the VoIP_Action. As hinted at by the sequence number 10, each policy can have multiple class-action associations. Just keep in mind where you will apply that policy, to ensure that class-action pairs are appropriate for where you intend to apply the policy, as described below.

4. Apply the policy

Access config-interface mode on the appropriate interface and apply the policy – always as an inbound policy. In other words, the policy is applied as packets enter the router inbound. This makes sense, correct? The purpose of PBR is to override normal destination-based routing. If you applied the policy outbound on an interface, it would be too late – routing decisions are already made.



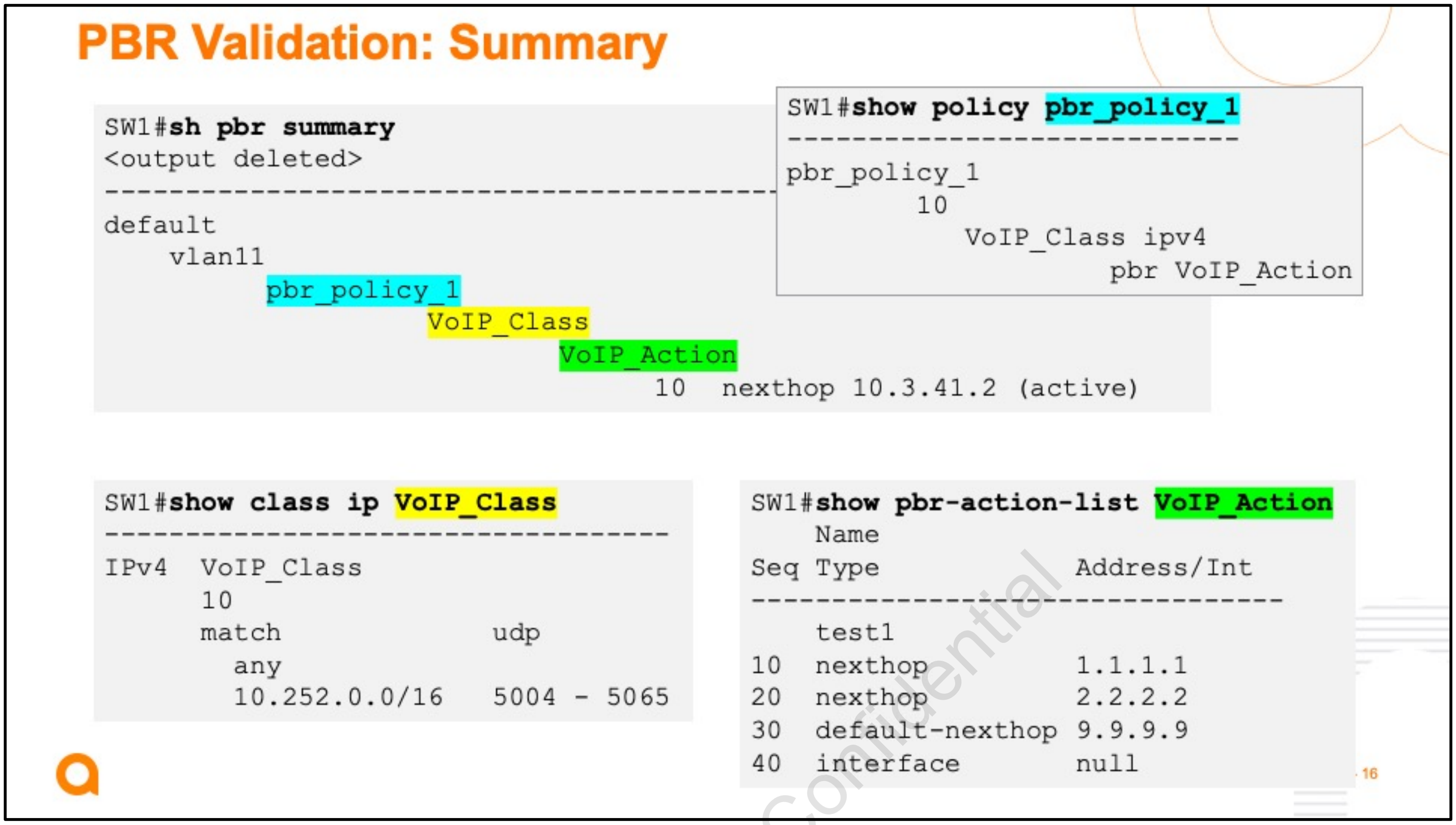
PBR Validation: Interface

```
SW1#sh pbr interface vlan11
<output deleted>
--------------------------------------------------------------------------------
default
    vlan11
          pbr_policy_1
                  VoIP_Class
                            VoIP_Action
                                  10  nexthop         10.3.41.2 (active)
                                  20  nexthop         10.3.42.3
                                  30  default_nexthop 10.3.1.2
                                  40  interface       null
```

MOD 1- 15

Use show pbr interface vlan11, for example, to see a nice overview of how all PBR pieces fit together in a kind of “configuration chain”. You see that pbr_policy_1 is applied to interface VLAN11. That policy will apply the PBR actions defined in VoIP_Action to all traffic in class VoIP_Class. And you can see the actions listed, and the top action is currently active – it’s a valid route, and so the other actions will not be used.

It can be helpful to compare this output to the appropriate portions of your running configuration, from the previous the figure.



PBR Validation: Summary

```
SW1#sh pbr summary
<output deleted>
------------------------------------------
default
    vlan11
          pbr_policy_1
                    VoIP_Class
                            VoIP_Action
                                  10   nexthop 10.3.41.2 (active)
```

```
SW1#show policy pbr_policy_1
----------------------------
pbr_policy_1
        10
           VoIP_Class ipv4
                    pbr VoIP_Action
```

```
SW1#show class ip VoIP_Class
-----------------------------------
IPv4  VoIP_Class
      10
      match                udp
        any
        10.252.0.0/16      5004 - 5065
```

```
SW1#show pbr-action-list VoIP_Action
    Name
Seq Type                Address/Int
---------------------------------
    test1
10  nexthop             1.1.1.1
20  nexthop             2.2.2.2
30  default-nexthop     9.9.9.9
40  interface           null
```

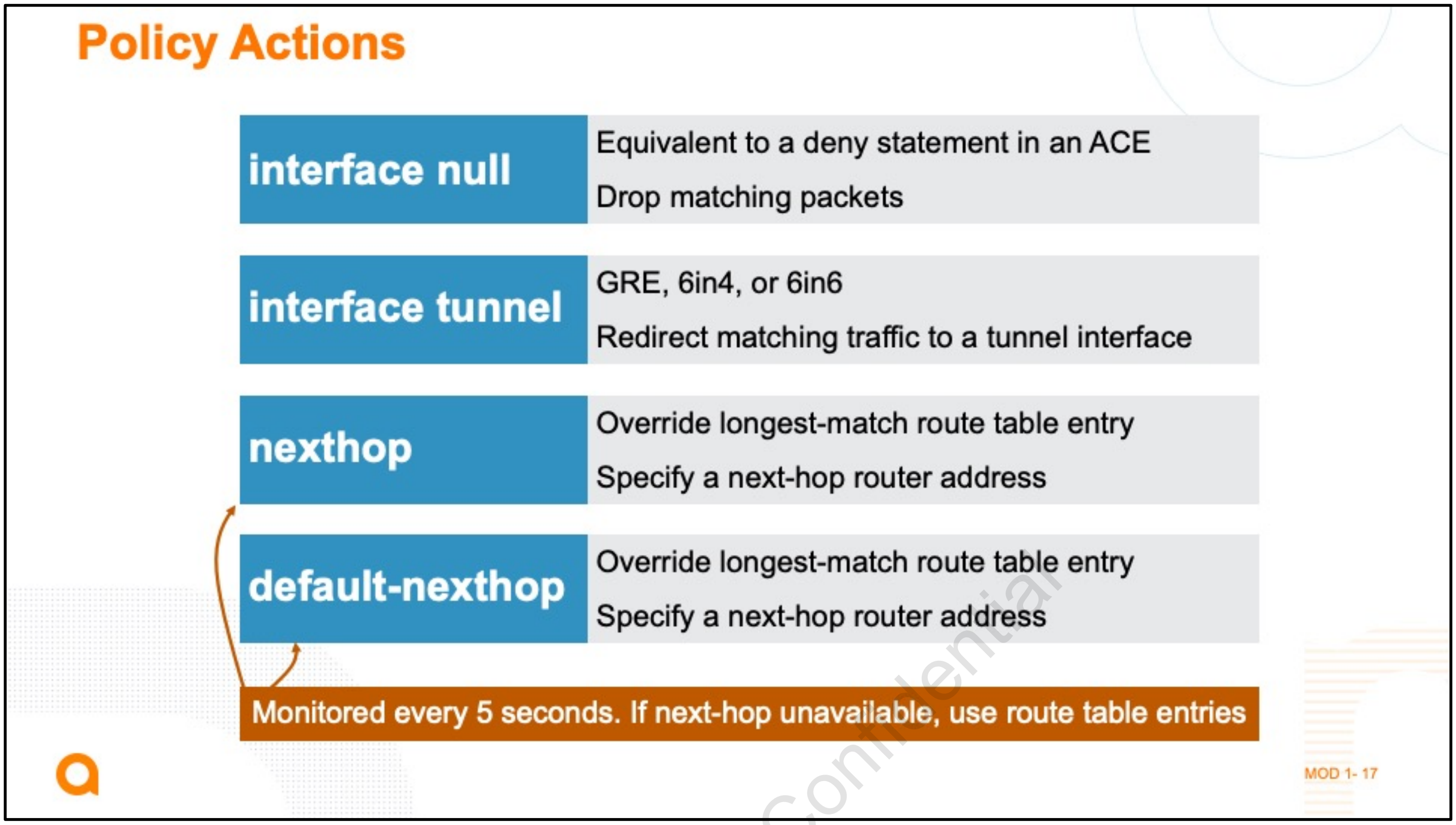
If you were not aware whether or how PBR is configured, show pbr summary can be helpful – it shows all PBR “configuration chains”. In the example switch SW1 only has one configuration chain. You now know the names of the policy, class, and action. Use this information to enter the appropriate commands to see more detail on each PBR component:

show policy pbr_policy_1

show class ip VoIP_Class

show pbr-action-list VoIP_Action

For more information, see the Aruba 10.xx IP Routing Guide.



Policy Actions

interface null	Equivalent to a deny statement in an ACE Drop matching packets
interface tunnel	GRE, 6in4, or 6in6 Redirect matching traffic to a tunnel interface
nexthop	Override longest-match route table entry Specify a next-hop router address
default-nexthop	Override longest-match route table entry Specify a next-hop router address

Monitored every 5 seconds. If next-hop unavailable, use route table entries

MOD 1- 17

For your reference, the figure summarizes the PBR policy actions you just learned about in the PBR configuration example.

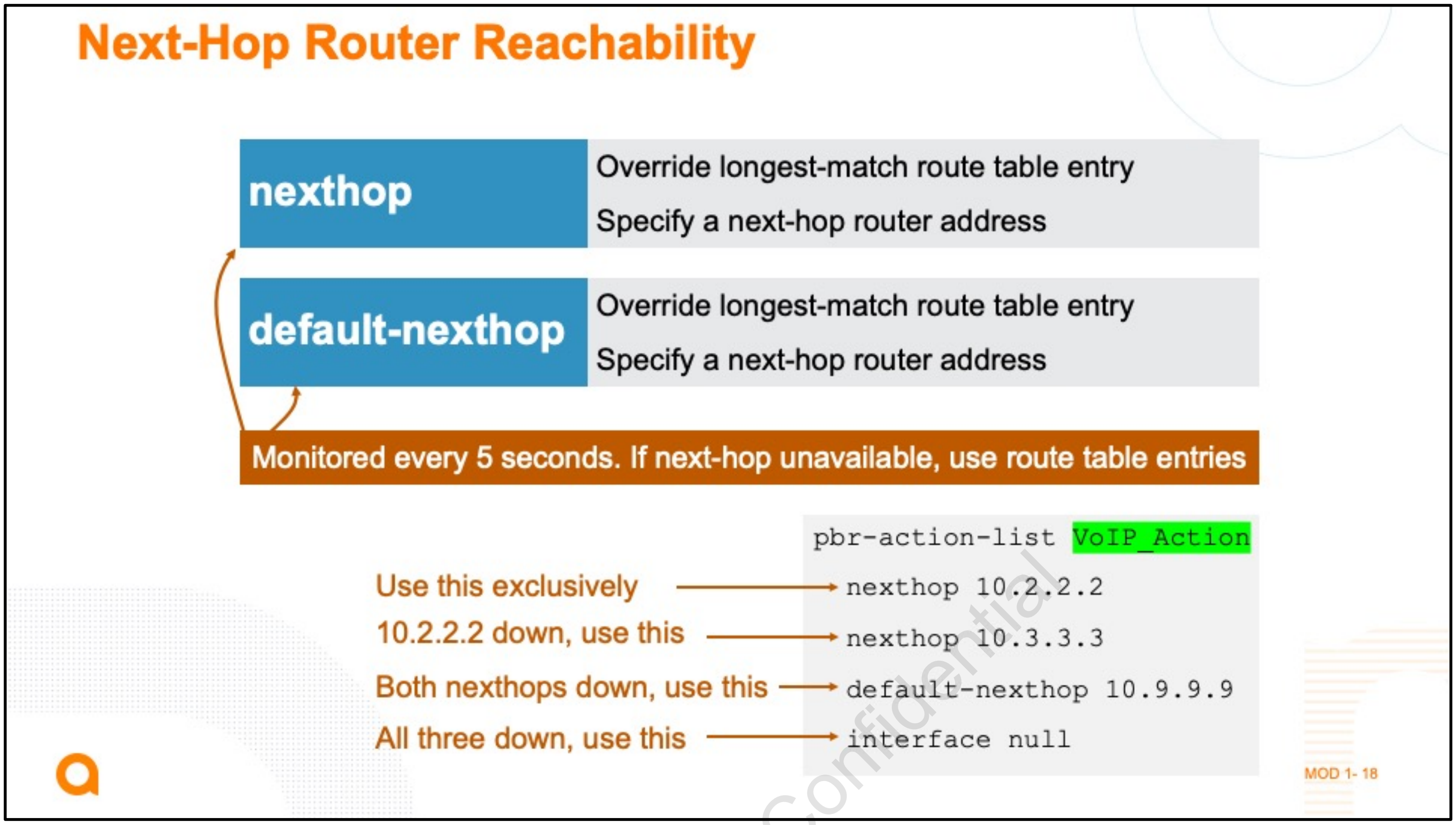
interface null: Packets that match the class criteria for that policy entry are dropped.

interface tunnel: Specify a GRE, 6in4 or 6in6 tunnel as the outbound interface for all matching packets. The tunnel must exist before configuring. Packets sent into the tunnel interface egress at the router at the endpoint of the tunnel. If the tunnel is misconfigured or down the traffic may be lost.

nexthop: Override the routing table's longest prefix match next-hop router with a next-hop address that you specify. If no such routing table entry exists for matching packets, (default or not) this action still affects matching packets.

default-nexthop: Specify a next-hop router for matching packets when there is no longest prefix match for those packets in the routing table. Such a default-nexthop overrides a system default route if already configured and also applies if there is no system default route.

Note: next-hop and default-nexthop facilitate routing matching packets that might not otherwise be routed - due to the absence of routing table entries. Unlike nexthop, default-nexthop only applies if there is no destination lookup match in the main routing table for matching packets. If the next-hop address for either command is not available, then the normal routing table entry (or entries are used).



Next-Hop Router Reachability

nexthop
Override longest-match route table entry
Specify a next-hop router address

default-nexthop
Override longest-match route table entry
Specify a next-hop router address

Monitored every 5 seconds. If next-hop unavailable, use route table entries

Use this exclusively →
10.2.2.2 down, use this →
Both nexthops down, use this →
All three down, use this →

```
pbr-action-list VoIP_Action
  nexthop 10.2.2.2
  nexthop 10.3.3.3
  default-nexthop 10.9.9.9
  interface null
```

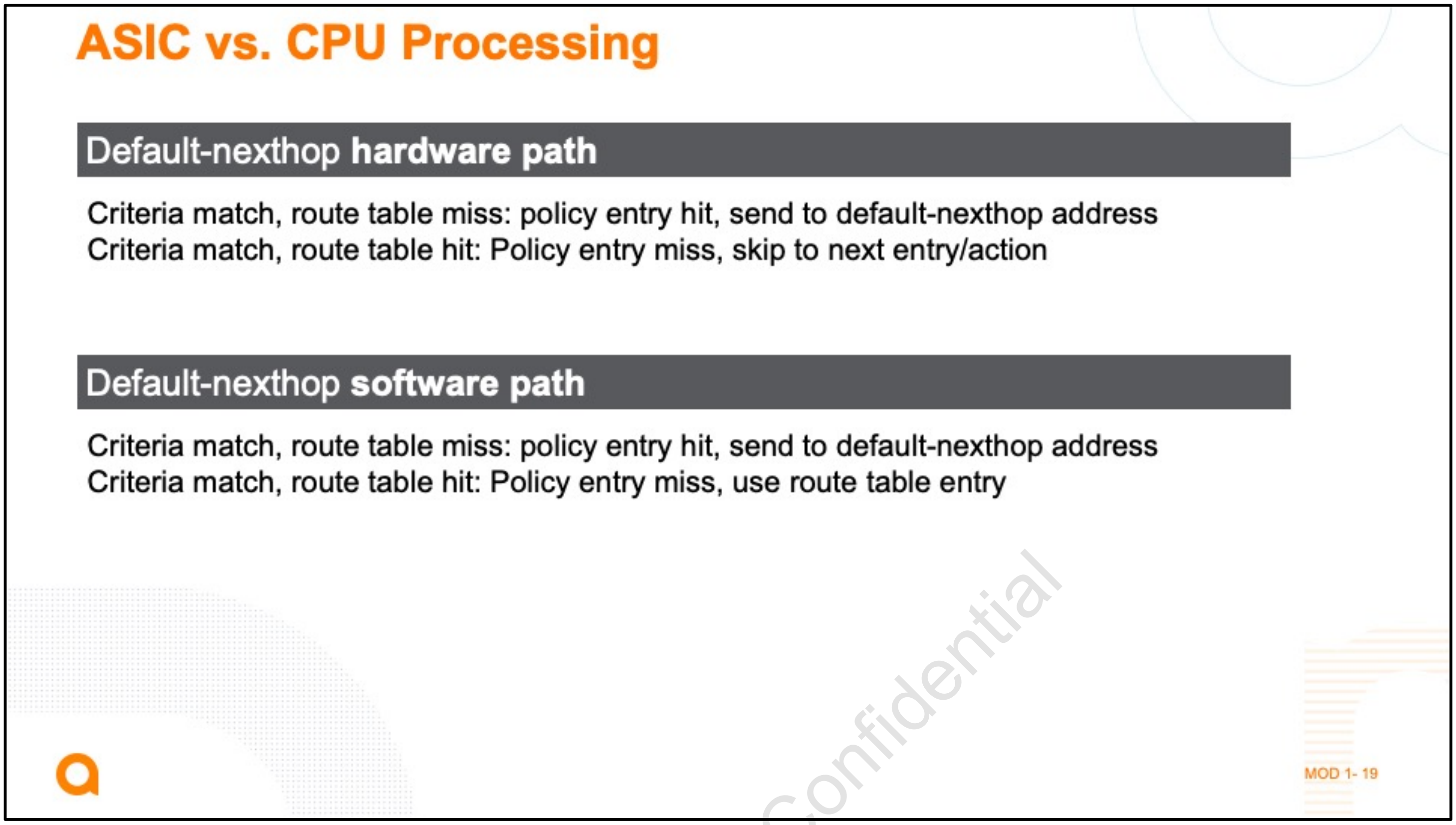
MOD 1- 18

PBR and next-hop router reachability

When you apply a PBR policy to an interface, all next-hop and default-nexthop entries in the associated action list are monitored every 5 seconds for reachability. If the next hop or default next hop becomes unavailable, the router falls back to using entries in the normal, destination-based route table.

The figure shows two reachable next-hops in an applied action list. The active entry is the one with the lowest sequence number – the one entered first, by default. This next hop is used exclusively unless that address becomes unavailable. Within 5 seconds, the router uses the next entry – 10.3.3.3. If both of those are down, then the router uses the default-nexthop. If all three are down, the interface null action ensures packets are merely thrown away. If you omit the interface null command, then the router falls back to using destination-based route table entries.

If you only have a single nexthop entry, the router falls back to destination-based routing when that path becomes unavailable. Once the next-hop IP address becomes available, it is once again used – within 5 seconds of course.



ASIC vs. CPU Processing

Default-nexthop **hardware path**

Criteria match, route table miss: policy entry hit, send to default-nexthop address
Criteria match, route table hit: Policy entry miss, skip to next entry/action

Default-nexthop **software path**

Criteria match, route table miss: policy entry hit, send to default-nexthop address
Criteria match, route table hit: Policy entry miss, use route table entry

MOD 1- 19

Traffic to be routed that cannot be handled by the switching ASIC includes IPv4 packets with IP options. This traffic is handled by the switch operating system kernel routing software. The characteristics of network usage determine the proportion of PBR traffic handled by software vs hardware path. Under normal conditions, software path PBR is likely to be extremely low – a fraction of one percent.

There is a difference between PBR hardware and software path when using the PBR default-nexthop action. The figure compares what happens when there is a policy match with an associated default-nexthop action for the hardware vs software path.

Hardware path

Condition: The policy criteria match, and there is no route table entry for the packet's destination IP address (route table "miss")

Result: Policy entry hit for that policy entry – send the packet to the specified default-nextop address

Condition: The policy criteria match, but there is a route table entry for the packet's destination (route table "hit")

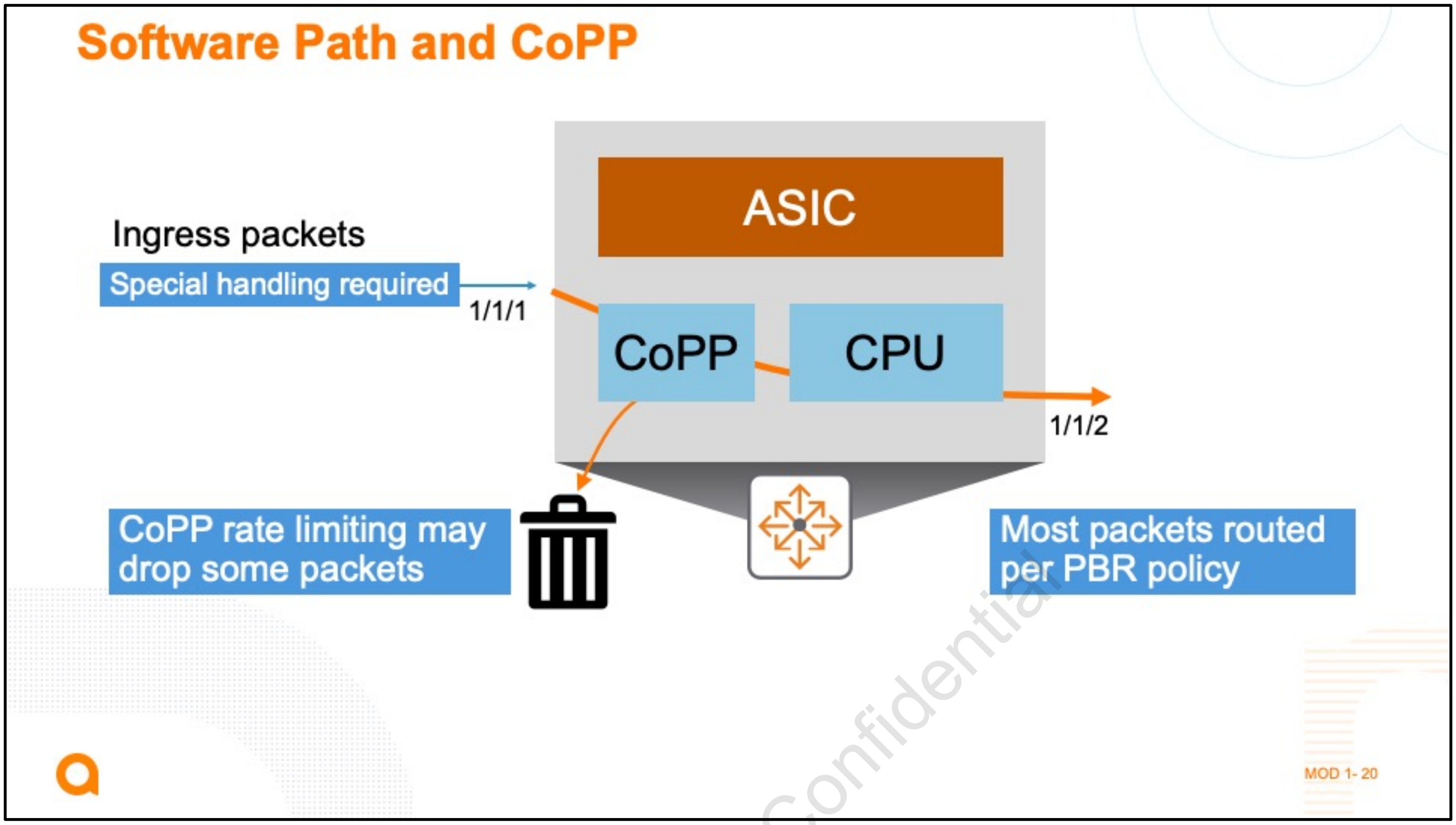
Result: Policy entry miss. Skip to the next entry/action in the policy.

Software path

Condition: IF the policy criteria matches AND there is a route table miss, THEN send forward to specified default-nexthop address.

Condition: IF the policy criteria matches AND there is a route table hit, THEN stop processing the policy and use normal destination-based routing.

This difference in behavior is due to a limitation in the software path matching and routing engine, relative to hardware.



Software Path and CoPP

Ingress packets

Special handling required

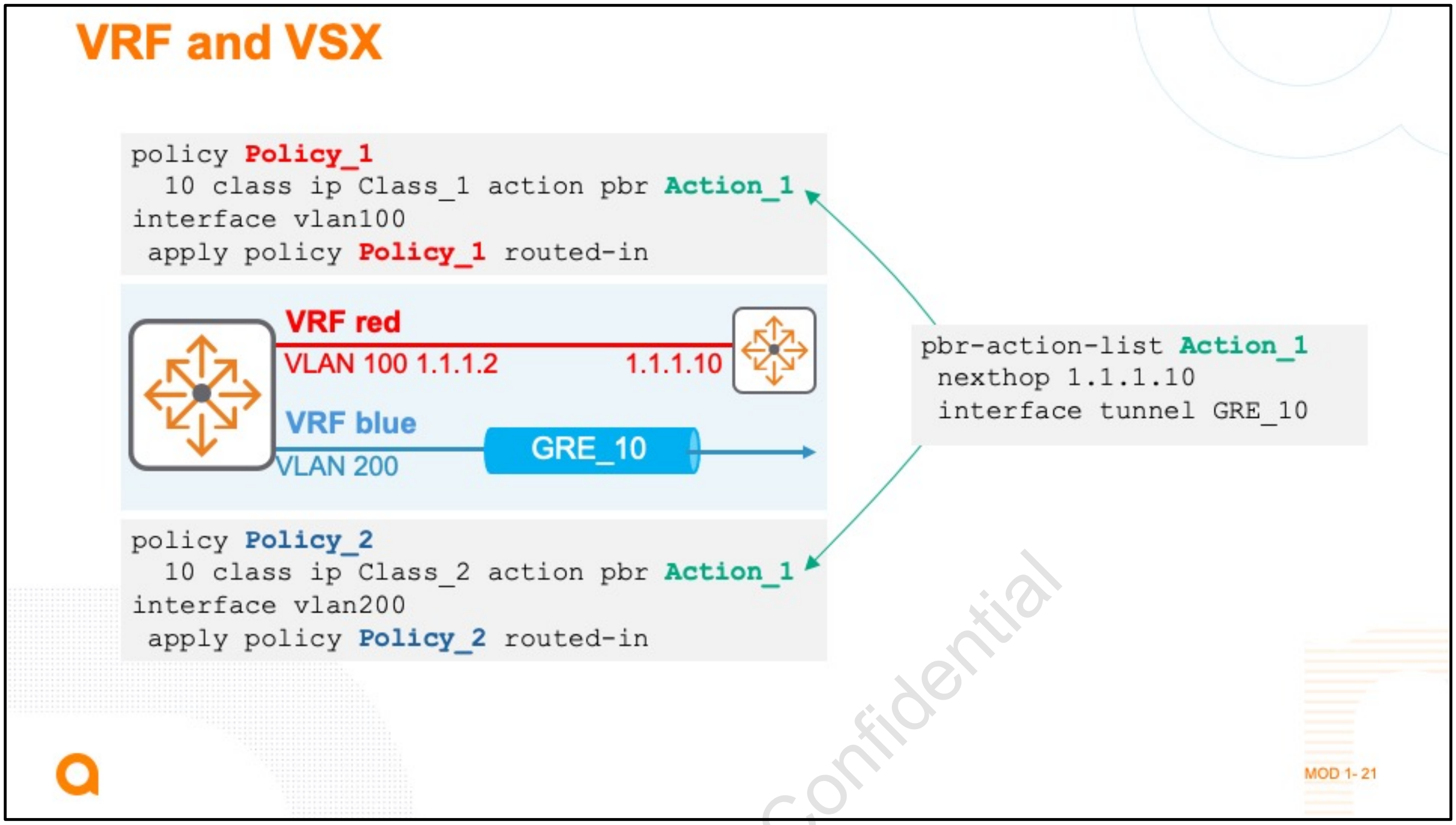
CoPP rate limiting may drop some packets

Most packets routed per PBR policy

Suppose that a packet arrives that requires special handling, and the router lacks a route table entry for its destination IP address – a route table “miss”. By default, an AOS-CX device sends these packets to the system CPU for special handling as needed – such as processing certain ICMP packets, for example.

This CPU routing has higher priority than a PBR next-hop action. Even though a policy is applied with an entry that matches the traffic and specifies a PBR next-hop action which is reachable, the traffic will still be routed to the system CPU – again, due to the absence of a reachable next-hop in the system route table. It will not be processes by the desired PBR hardware path. It will be properly routed by PBR software path, but it will also be rate limited by the Control Plane Policing(CoPP) feature. Loss can occur at higher traffic rates.

Important: The workaround for this issue is to create a default next-hop route in the system with a reachable next-hop router/host. This will result in a route hit, or reachable next-hop detected for the traffic with no further need to route traffic to the CPU. The PBR hardware path next-hop action will then occur, as desired.



VRF and VSX

```
policy Policy_1
  10 class ip Class_1 action pbr Action_1
interface vlan100
 apply policy Policy_1 routed-in
```

```
pbr-action-list Action_1
 nexthop 1.1.1.10
 interface tunnel GRE_10
```

```
policy Policy_2
  10 class ip Class_2 action pbr Action_1
interface vlan200
 apply policy Policy_2 routed-in
```

MOD 1- 21

You can apply PBR policies to interfaces in any VRF, and PBR action lists can be used in different policies across different VRFs. The effect of applying the same Policies/PBR action lists across different VRFs depends on the IP networks and interfaces configured in the different VRFs.

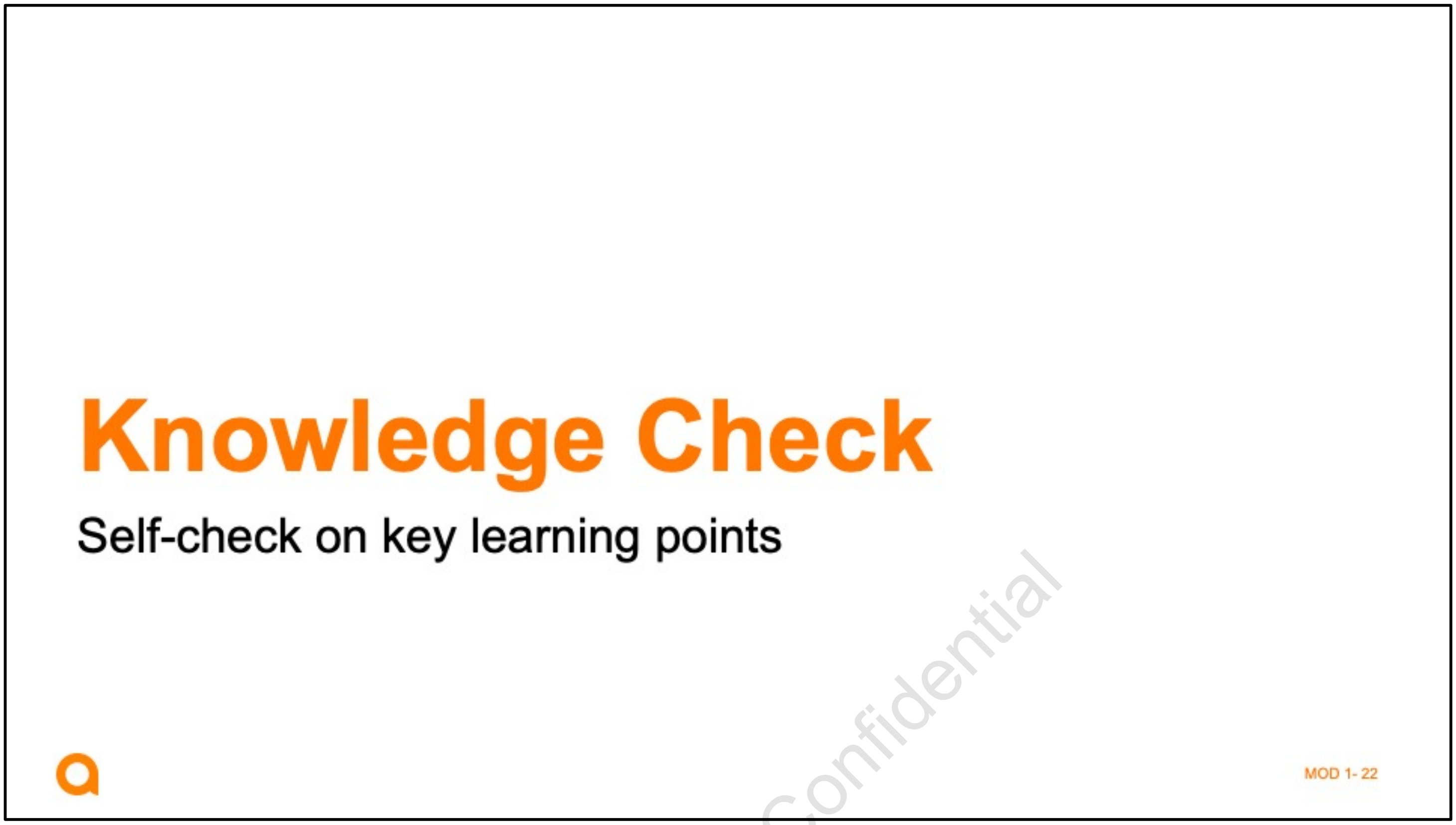
The figure shows an action list with actions nexthop 1.1.1.10 and interface tunnel gre_10, used as the PBR action parameter for an entry in policy_1 and also in policy_2. If you apply policy_1 to an interface in VRF 'red', which has an interface in subnet 1.1.1.0/24 but no GRE tunnel named 'gre_10', then only the 'nexthop' action is relevant to VRF 'red'. Suppose policy_2 contains the same action list. It is applied to an interface in VRF 'blue', which lacks the 1.1.1.0/24 subnet configured but does have a tunnel named 'gre_10'. In that case only that interface tunnel action will apply in that VRF.

Note: It is possible to configure the same subnet in different VRFs, however named tunnel interfaces can only exist in one, so in the example of a common action list, the 'next-hop' action could be relevant to both VRFs, but the 'interface tunnel' action may only be relevant to one. If VRFs are part of the router configuration, be mindful of them when creating and applying policies with PBR action lists and their entries. VRF Route Leaking is not supported in the current release of PBR (10.4).

PBR software path, VSX, and VRRP

Due to the dynamic nature of the VSX and VRRP protocols, applying a policy with PBR to VSX or VRRP Layer 3 kernel interfaces and then making further changes to those protocols may not result in expected behavior.

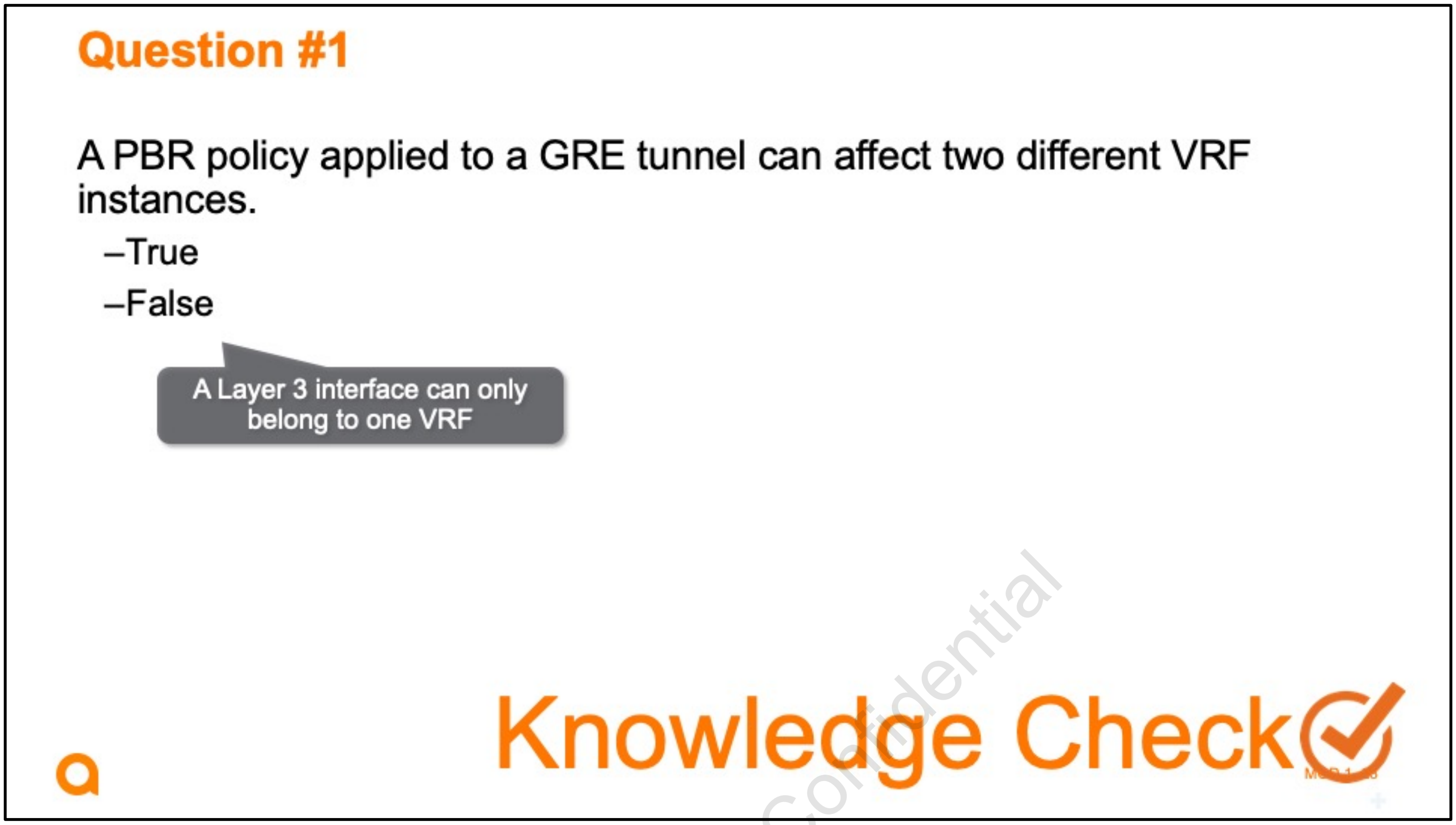
Apply policies with PBR to VSX and VRRP Layer 3 interfaces after all changes have been made. If further changes are necessary the policy should be removed first and reapplied when configuration updates are complete.



Knowledge Check

Self-check on key learning points

MOD 1- 22



Question #1

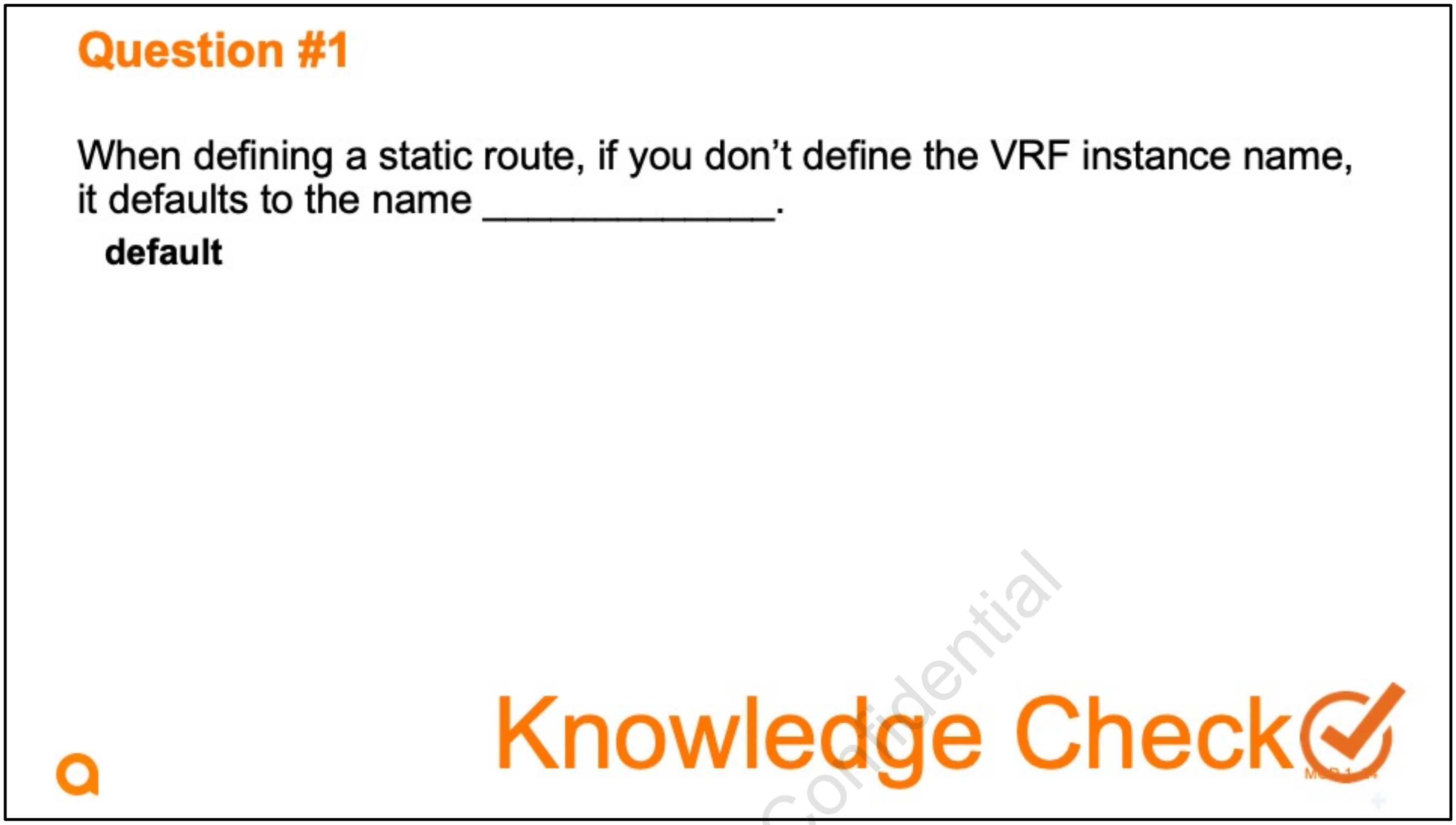
A PBR policy applied to a GRE tunnel can affect two different VRF instances.

–True

–False

A Layer 3 interface can only belong to one VRF

Knowledge Check

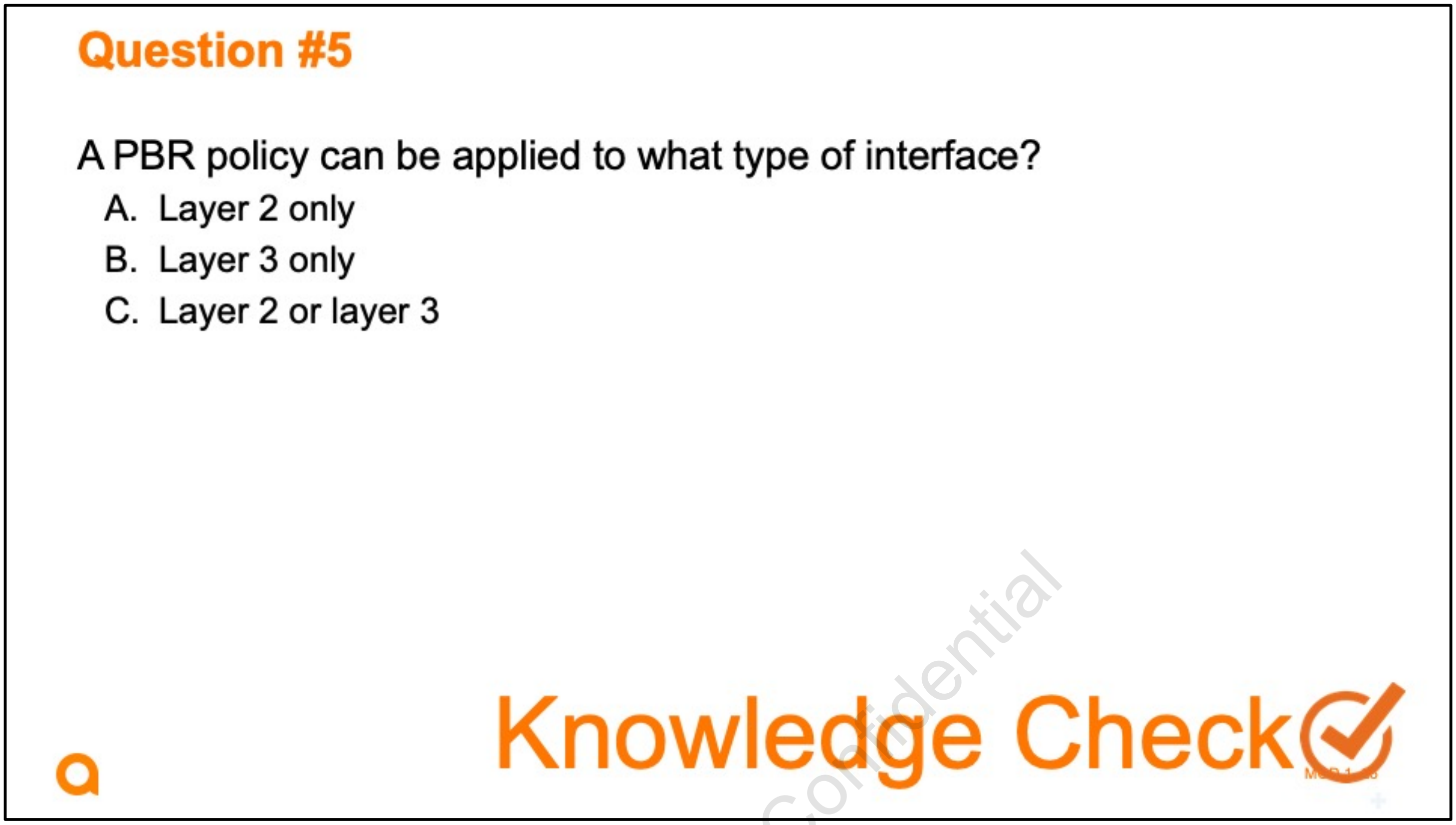


Question #1

When defining a static route, if you don't define the VRF instance name, it defaults to the name _____________.

default

Knowledge Check



Question #5

A PBR policy can be applied to what type of interface?

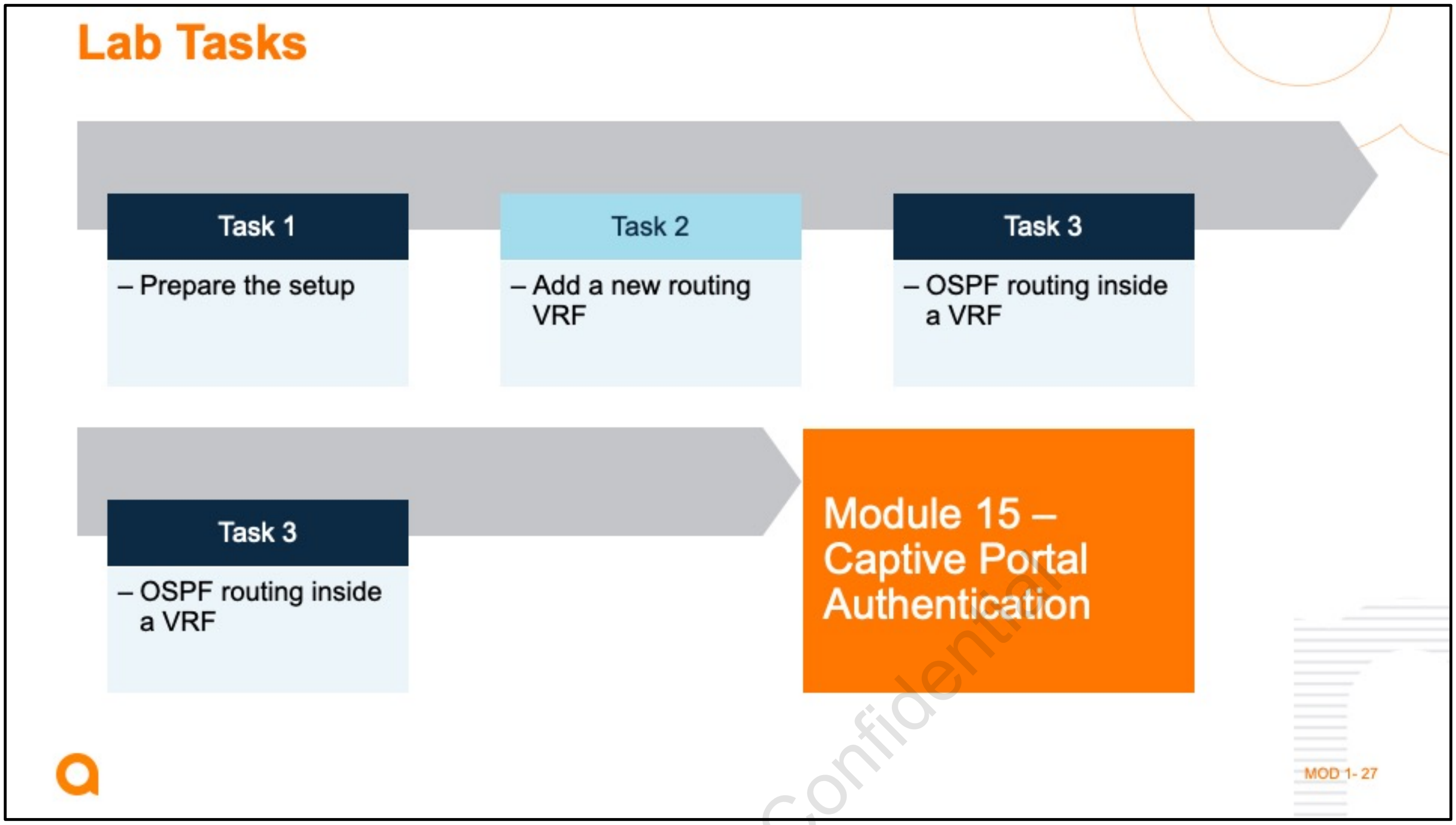
A. Layer 2 only

B. Layer 3 only

C. Layer 2 or layer 3

Knowledge Check

Lab Activity



Lab Tasks

Task 1

– Prepare the setup

Task 2

– Add a new routing VRF

Task 3

– OSPF routing inside a VRF

Task 3

– OSPF routing inside a VRF

Module 15 – Captive Portal Authentication

MOD 1- 27

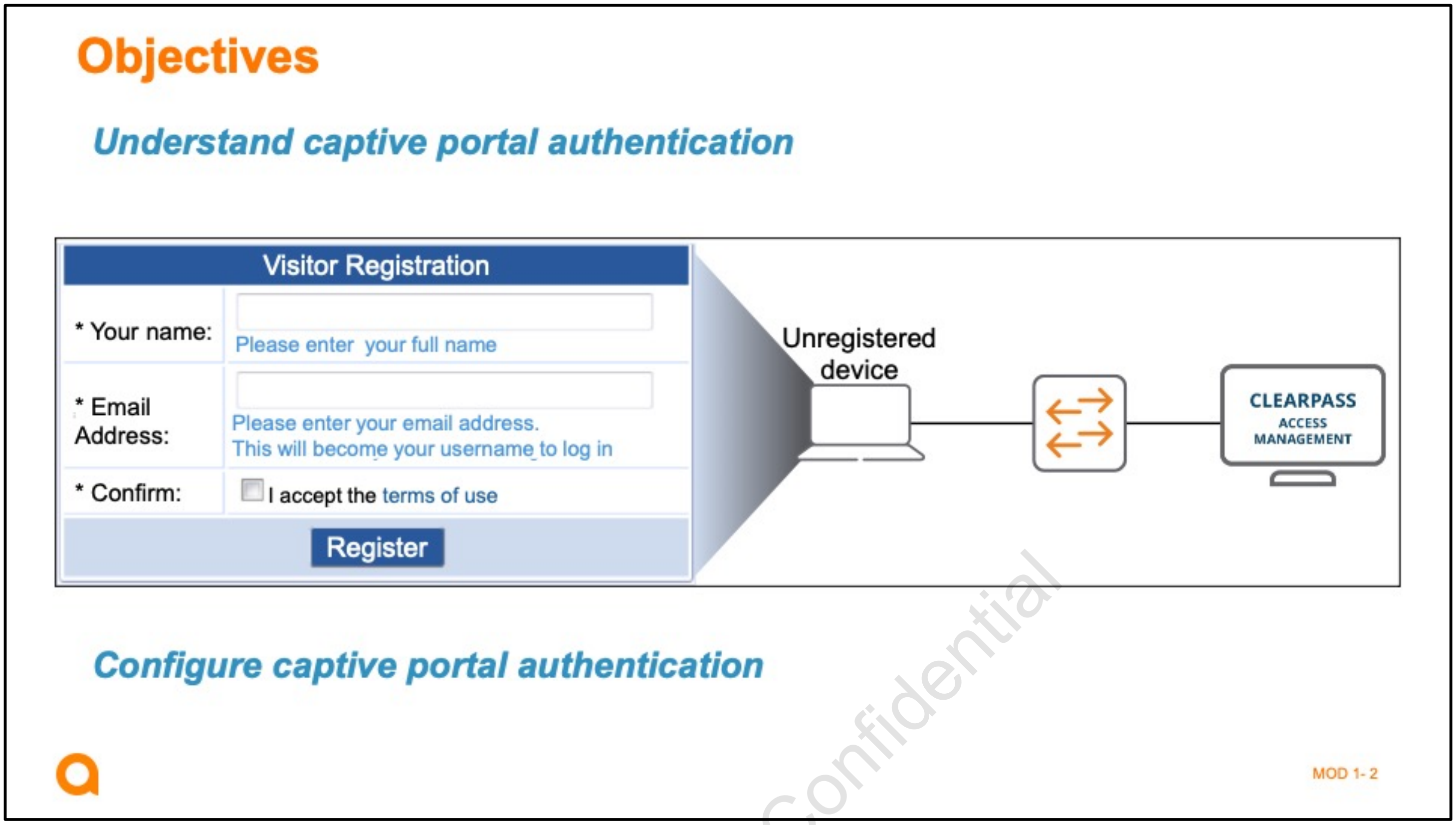
The figure provides a brief review of lab tasks. Please see your lab guide for details.

Rev # 20.21

Captive Portal Authentication

Implementing ArubaOS-CX Switching

aruba
a Hewlett Packard Enterprise company



Objectives

Understand captive portal authentication

Visitor Registration

* Your name:	Please enter your full name
* Email Address:	Please enter your email address. This will become your username to log in
* Confirm:	□ I accept the terms of use

Register

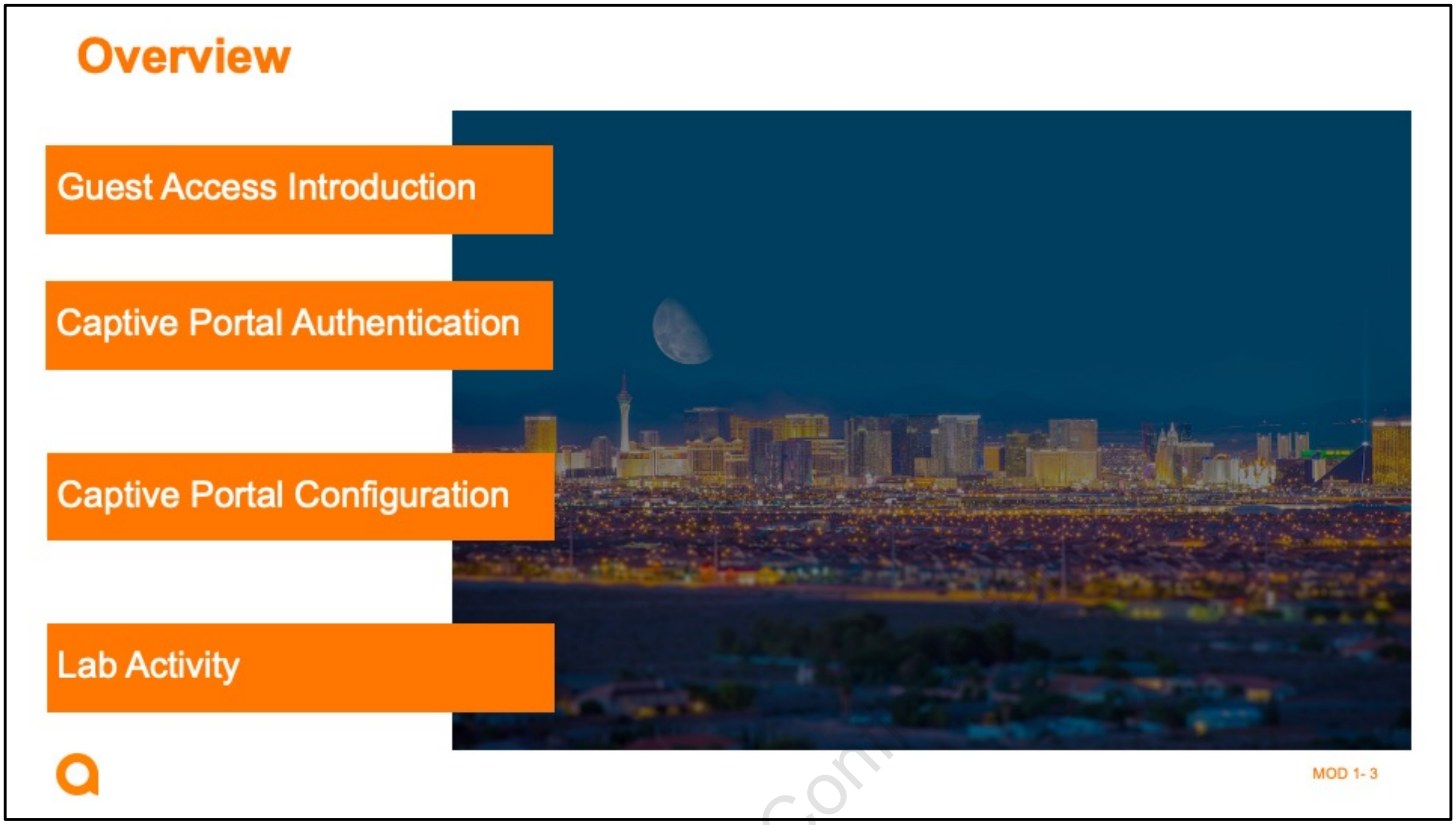
Configure captive portal authentication

MOD 1- 2

After completing this module, you should be able to:

Understand the process AOS-CX switches use for captive portal authentication

Configure captive portal authentication on AOS-CX switches to integrate them with an Aruba ClearPass guest and BYOD solution



Overview

Guest Access Introduction

Captive Portal Authentication

Captive Portal Configuration

Lab Activity

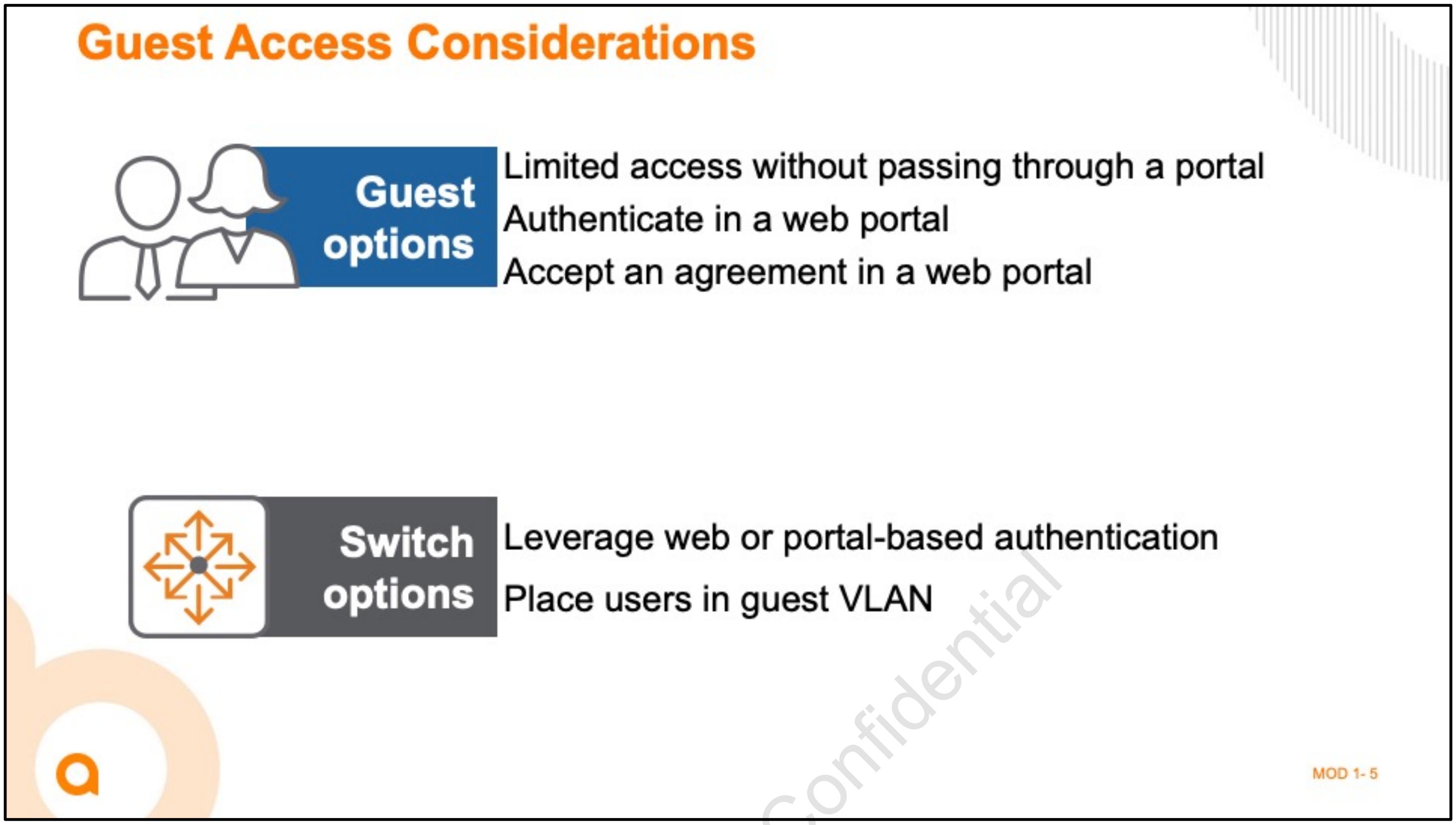
MOD 1- 3

There are situations where a company wants to control access to a network, perhaps for guest access, but using 802.1X or MAC-Auth is not practical. This module teaches you how to set up AOS-CX switches to participate in guest and BYOD access solutions by implementing a captive portal solution, including configuration, and a lab activity

Guest Access Introduction

aruba
a Hewlett Packard
Enterprise company

This section introduces guest access.



Guest Access Considerations

Guest options

- Limited access without passing through a portal
- Authenticate in a web portal
- Accept an agreement in a web portal

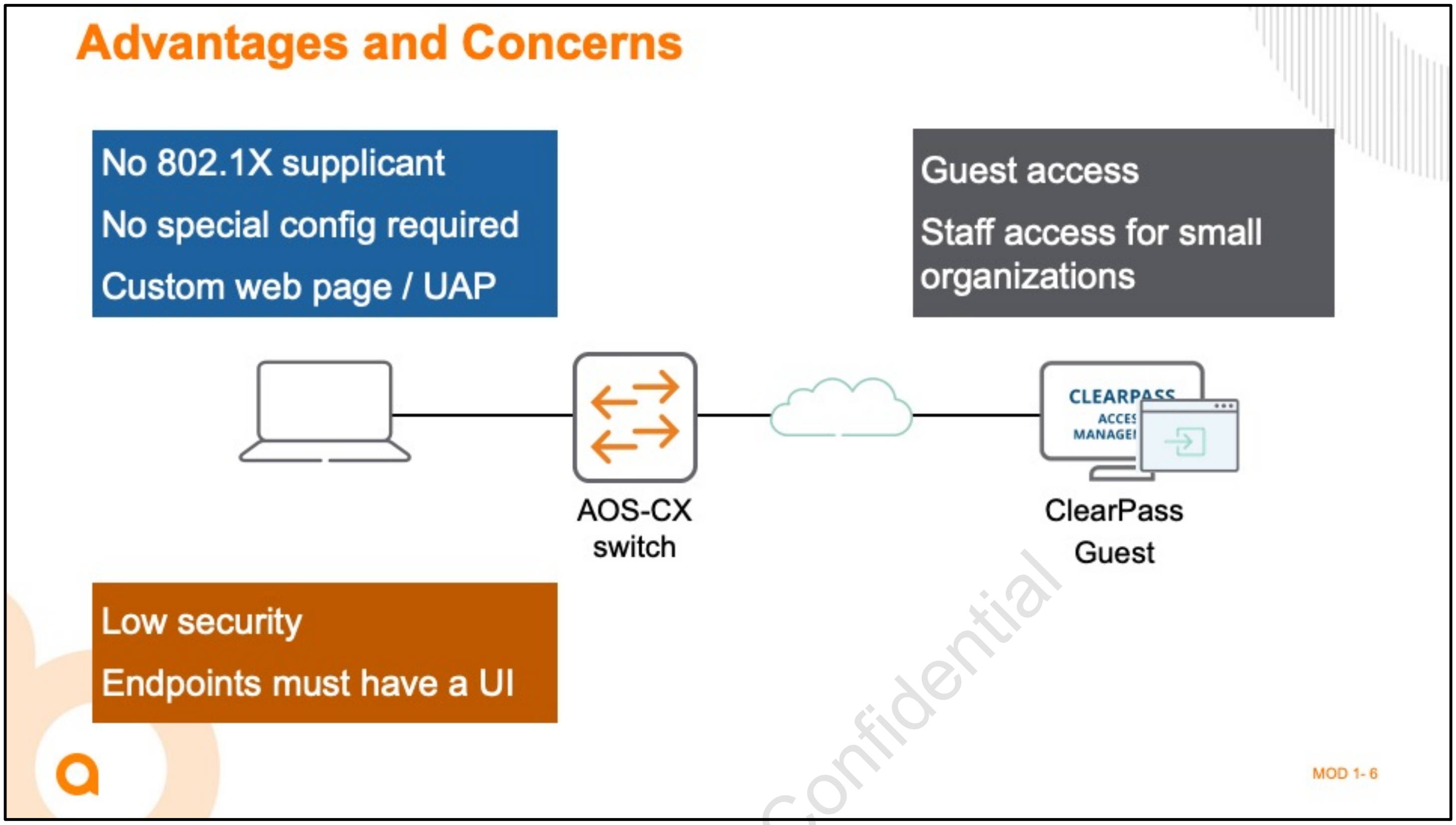
Switch options

- Leverage web or portal-based authentication
- Place users in guest VLAN

MOD 1- 5

Different companies have different expectations and needs for their guest access solution. Some companies might allow guests ^limited access without any authentication. However, most companies want guests to log in through a Web portal of some kind. Even if they do not care whether the guests authenticate, they want the guests to accept an agreement for legal purposes.

The company’s expectations will dictate whether you need to configure AOS-CX switches at the access layer to enforce or participate in a Web-based authentication solution, or whether you can simply place guests in a guest VLAN. You will focus on the Web-based authentication—also called portal authentication—solutions for most of this module. At the end, however, you will learn how you can enforce authentication for employees, but allow limited access to guests without forcing the guests to log in.



Advantages and Concerns

No 802.1X supplicant

No special config required

Custom web page / UAP

Guest access

Staff access for small organizations

Low security

Endpoints must have a UI

Web-based authentication eases user login – no special device requirements, no user expertise. Users simply open a Web browser and are directed to a login page. They might enter credentials, or simply accept an Acceptable Use Policy (AUP) agreement to log in.

The simplicity and lower security make this perfect for guest access. However, some small organizations may use web-based authentication for employees.

Web-Auth Advantages

Web-Auth also lets companies control what users see before they receive network access. The company can present a customizable webpage that shows the guest information about the company providing the complimentary access. The company can also force the user to consent to terms and conditions, which can be a legal requirement in many regions.

Environments with external users who may benefit from Web-Auth include:

Hospitals

Universities

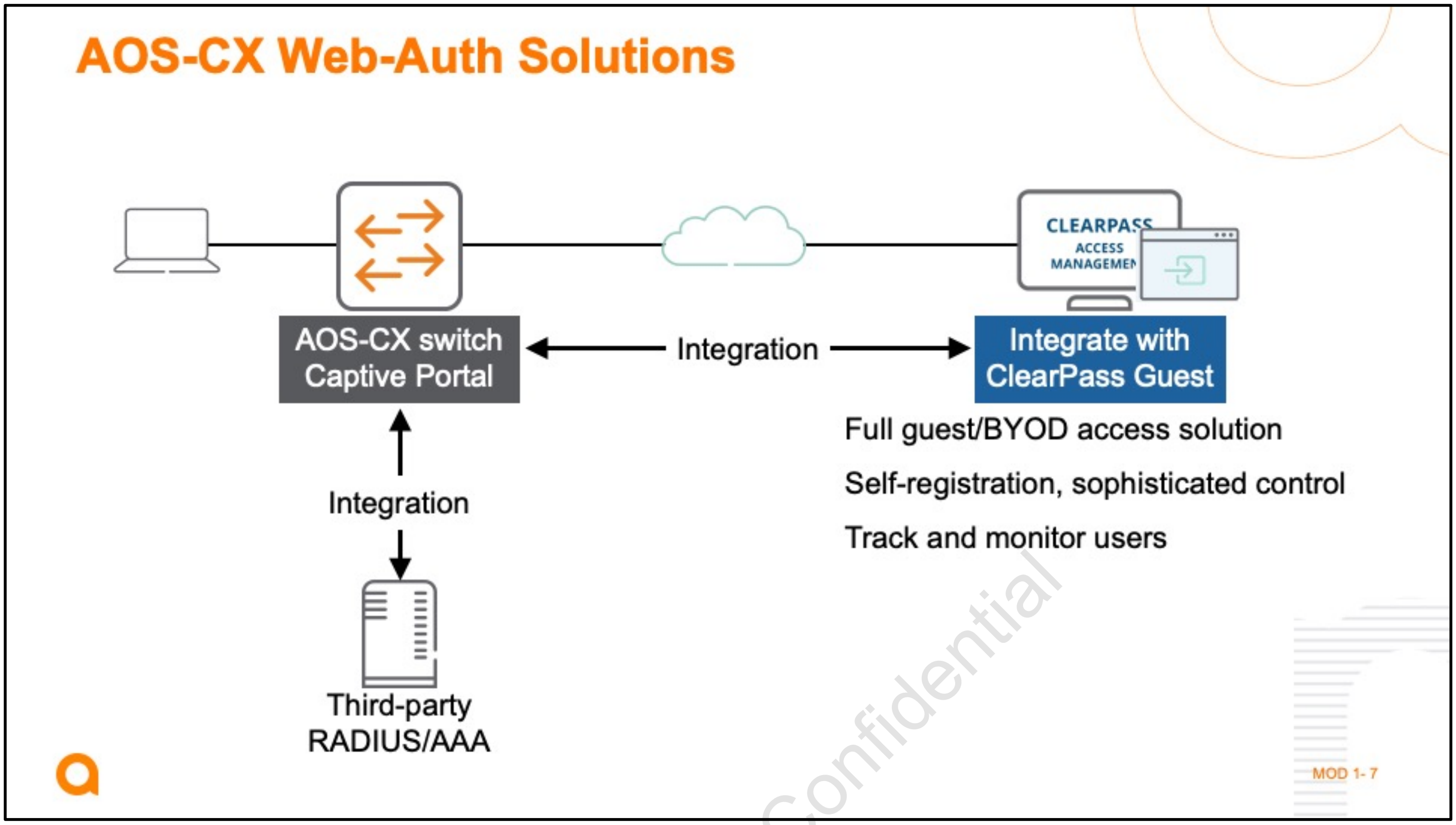
Libraries, hotels, and other businesses that provide courtesy Ethernet access

Web-Auth Disadvantages

Web-Auth does not provide as rigorous security as 802.1X. Users have a degree of network access before they authenticate. The credentials and authentication protocol might not be as secure as with 802.1X; however, you can improve the security by using HTTPS rather than HTTP.

Because Web-Auth requires interaction with the user, you cannot use it to authenticate devices without a user interface.

You should also note that the local Web-Auth features are quite limited. However, the switches can integrate with Aruba ClearPass, which is a sophisticated guest and BYOD solution.



AOS-CX Web-Auth Solutions

Full guest/BYOD access solution

Self-registration, sophisticated control

Track and monitor users

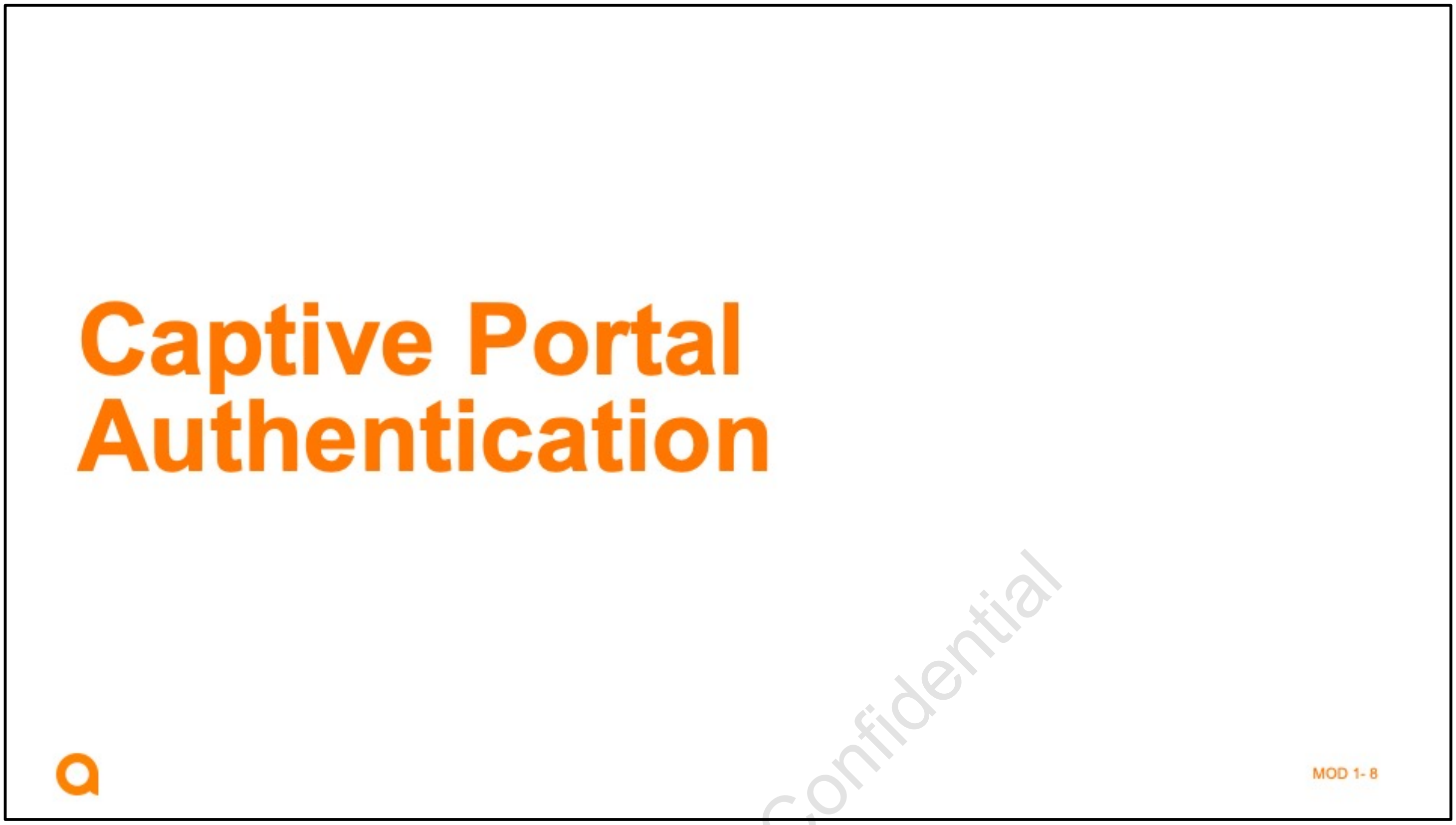
MOD 1- 7

In the past, Web-based authentication was as simple as a login page and perhaps a backend authentication server. However, these solutions have evolved. A solution such as ClearPass provides a Web portal as part of its complete user access management solution. Such a solution helps users log in more easily with features such as self-registration. It also helps companies track and monitor users, and it controls their access in sophisticated ways.

AOS-CX switches currently support one very capable web-based authentication method – captive portal. You get a full guest and BYOD access management solution, AOS-CX switches can use captive portal authentication to integrate with Aruba ClearPass. Aruba ClearPass provides the RADIUS server and portal server; the Policy Manager and Guest components are required for this solution. ClearPass OnGuard is required if companies want the endpoint integrity features. Note that AOS-CX switches can also integrate with other third-party AAA guest solutions.

Currently AOS-CX switches do not support a local Web-Auth, where the switches provide the web server and authentication component locally on the switch.

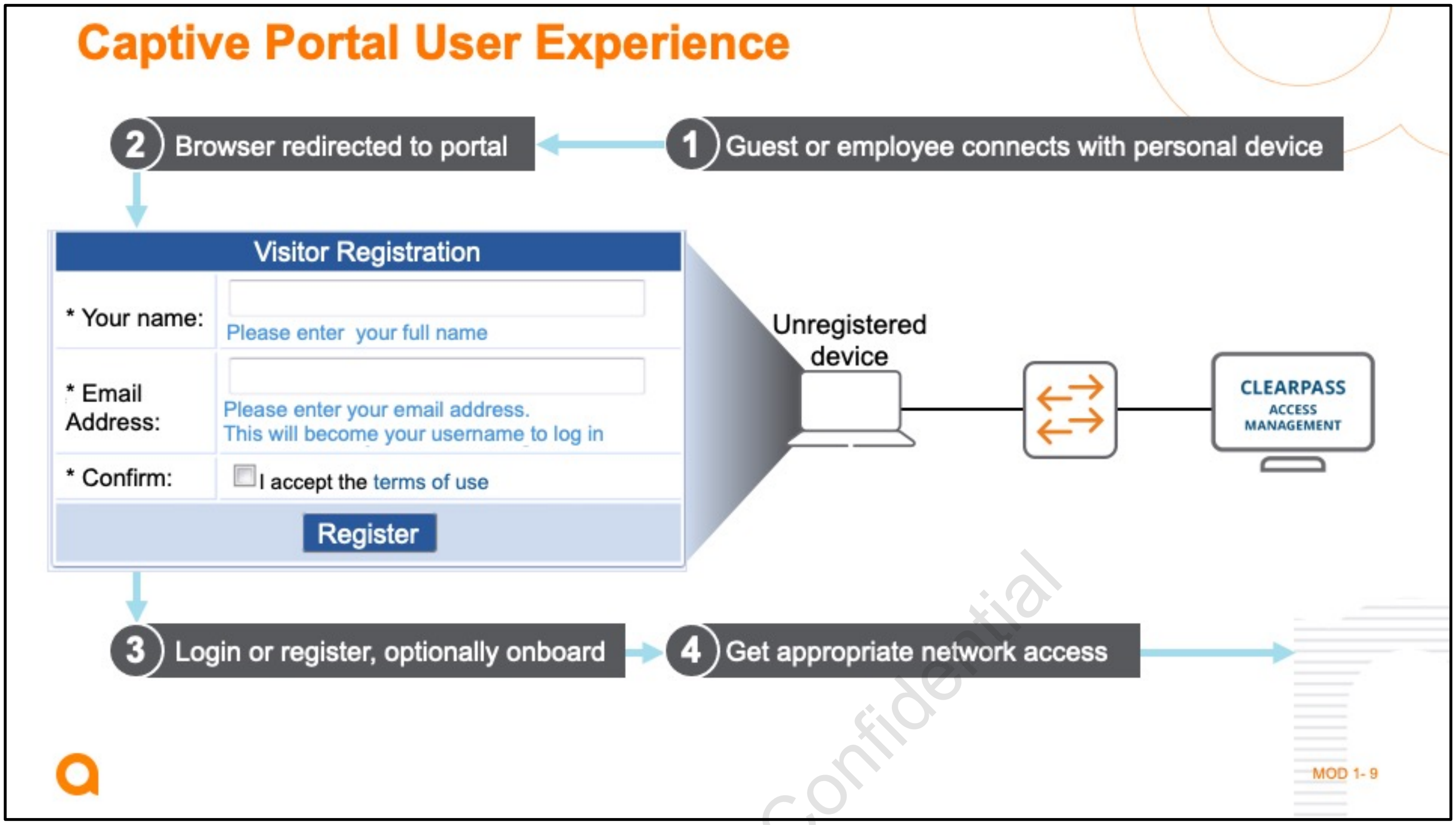
Note: AOS-CX switches can also integrate with HPE IMC’s BYOD solution using the BYOD redirect feature. Or it can use MAC Auth with Failure Redirect (MAFR) to redirect users who fail MAC authentication to register with a server that then adds their MAC address to the authorized list. These features are intended for similar use cases as the captive portal feature, and the use of each feature is mutually exclusive. BYOD redirect and MAFR are not covered in this training.



Captive Portal Authentication

MOD 1- 8

You will now learn how the captive portal process works with AOS-CX switches.



Captive Portal User Experience

1 Guest or employee connects with personal device

2 Browser redirected to portal

Visitor Registration

* Your name:	Please enter your full name
* Email Address:	Please enter your email address. This will become your username to log in
* Confirm:	□ I accept the terms of use

Register

3 Login or register, optionally onboard

4 Get appropriate network access

MOD 1- 9

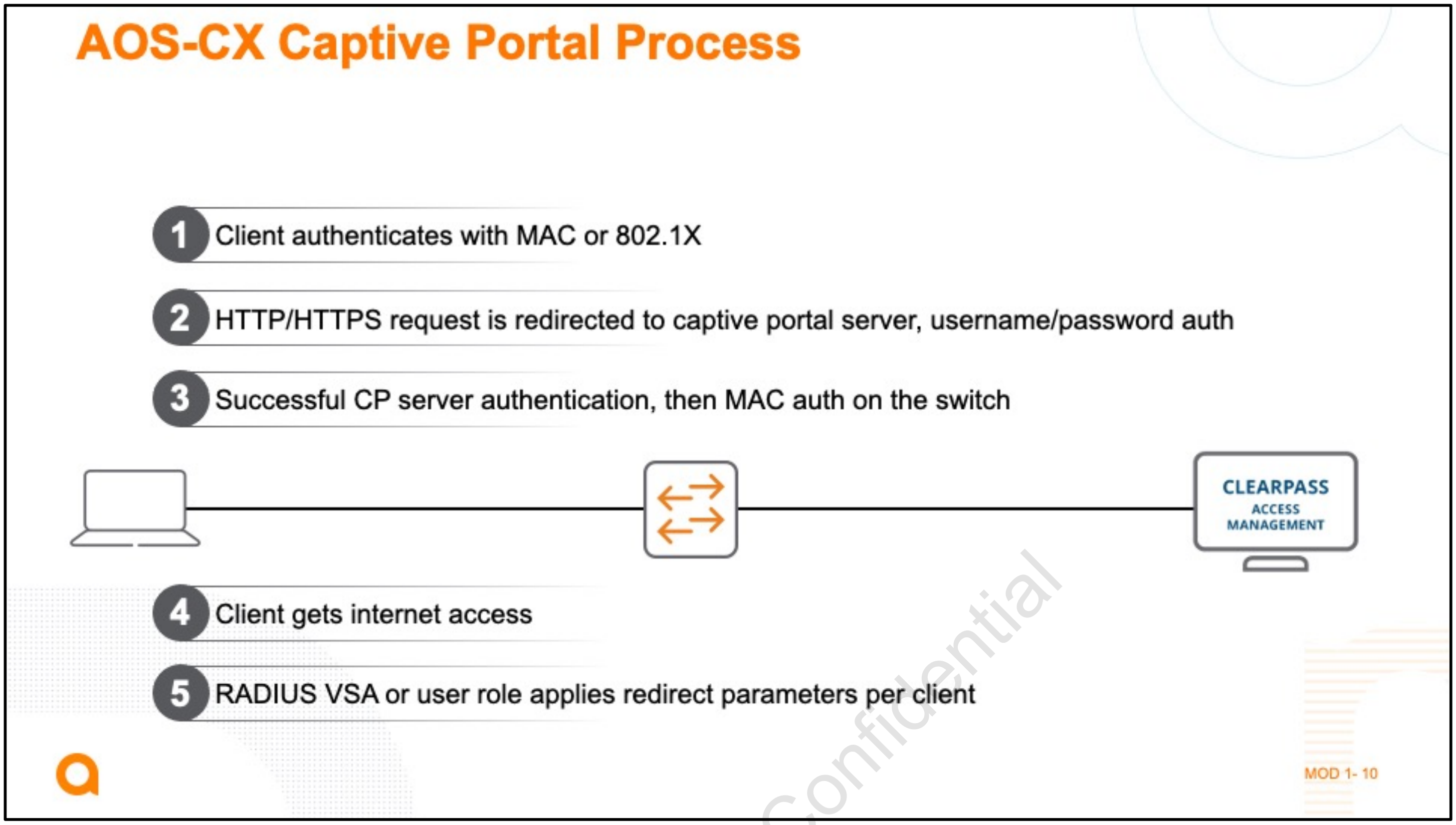
The captive portal authentication feature requires close teamwork between the AOS-CX switch and Aruba ClearPass. First let’s look at the user experience.

A user connects a new device an AOS-CX switch port, or associates with an Aruba Access Point (AP). The user powers on the device and opens the Web browser. The user sees a portal page, with options for logging in as an employee or a guest. At this point, the user device is locked down – they cannot browse to other pages or receive access to other network services.

Employees must enter valid employee credentials, perhaps stored in on Windows Active Directory. Guests use guest credentials to sign in, or they might be allowed to create their own account and log in with that. The options are controlled by the ClearPass Guest solution, which you can explore during the lab.

After logging in or registering for an account, the user is permitted network access, customized based on their identity and possibly other factors such as the health of their device. For example, an employee might receive different access rights than a guest. And a device that complies with security policies will receive more access than a potentially insecure device.

If the user connects with the same device later or in a different location, the device is provided the same access rights transparently.



AOS-CX Captive Portal Process

1 Client authenticates with MAC or 802.1X

2 HTTP/HTTPS request is redirected to captive portal server, username/password auth

3 Successful CP server authentication, then MAC auth on the switch

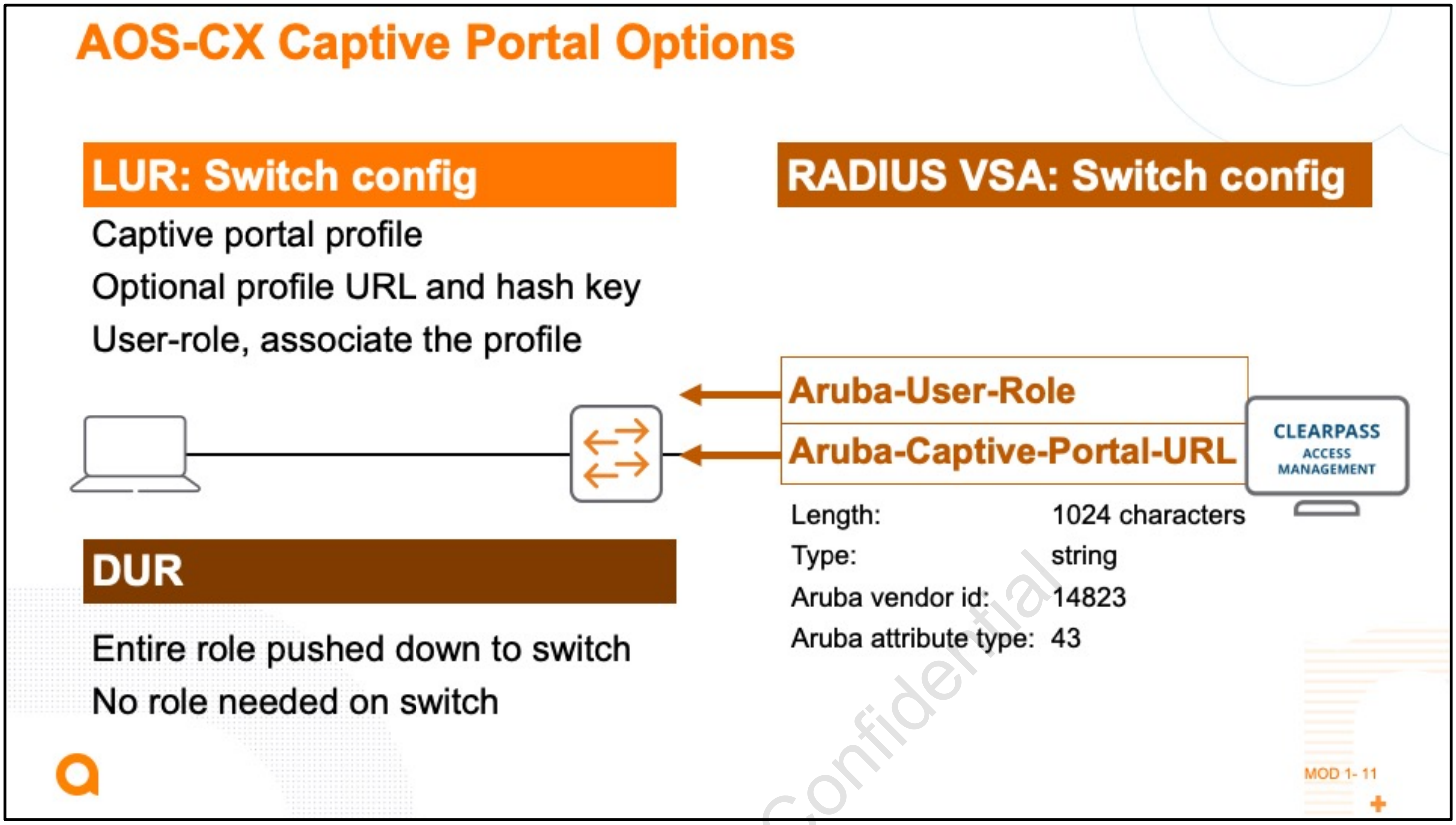
4 Client gets internet access

5 RADIUS VSA or user role applies redirect parameters per client

MOD 1- 10

The captive portal solution provides user access based on http/https redirection. First the client does MAC or 802.1X authentication. Then the http/https request is redirected to the captive portal (CP) server, which does username/password authentication. After this succeeds, the client does MAC authentication on the switch and gains internet access.

This feature is supported only for the clients getting authenticated through RADIUS server. This feature is client based. The redirect parameters per client can be applied user roles (local or downloadable) or RADIUS VSAs.



AOS-CX Captive Portal Options

LUR: Switch config

Captive portal profile

Optional profile URL and hash key

User-role, associate the profile

RADIUS VSA: Switch config

Length:	1024 characters
Type:	string
Aruba vendor id:	14823
Aruba attribute type:	43

DUR

Entire role pushed down to switch

No role needed on switch

| If you are implementing Local User Roles (LUR), the configuration steps on the switch includes:

Configure the captive portal profile.

Configure the URL and hash key(optional) in the captive portal profile.

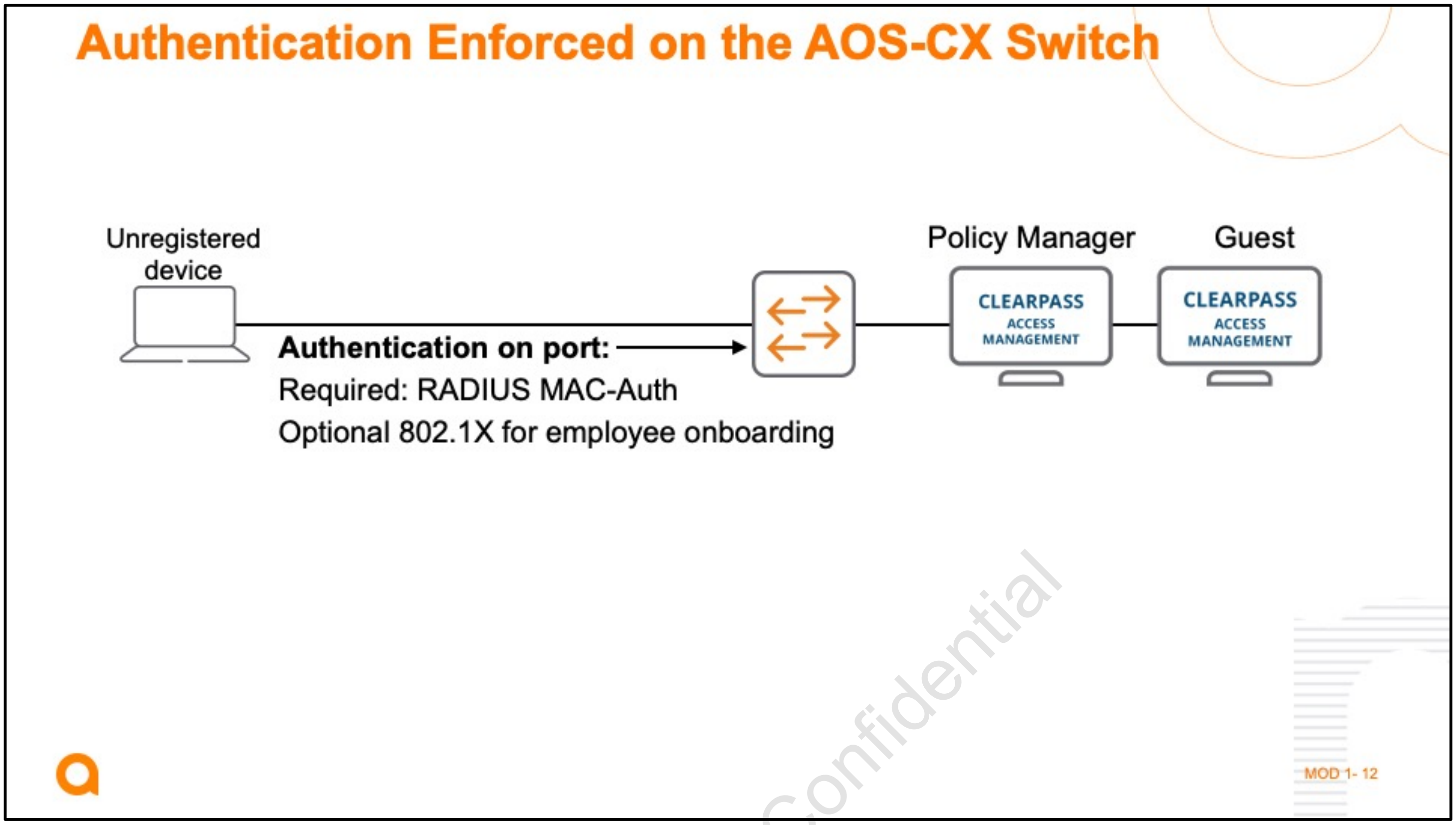
Configure user-role and associate the earlier configured captive portal profile.

|When using RADIUS VSA, the "Aruba-user-role" attribute is sent from RADIUS server in the authentication response packet. This role must be pre-configured on the switch with captive portal profile name. If the role is not in the switch, authentication fails. The following Aruba VSA's are in the authentication response for a client from RADIUS server:

Aruba-Captive-Portal-URL : This will have the captive portal profile URL information. A default role will be considered on the switch and gets applied to the client when this VSA is sent.

- Length: 1024 characters
- Type: string
- Aruba vendor ID: 14823
- Aruba attribute type: 43

| When implementing DUR, the captive portal profile and URL can be configured through the DUR for a client, along with the policies, and the entire role is pushed down to the switch. Through this process, the role doesn't need to exist on the switch: instead, it is centrally maintained on the ClearPass server.

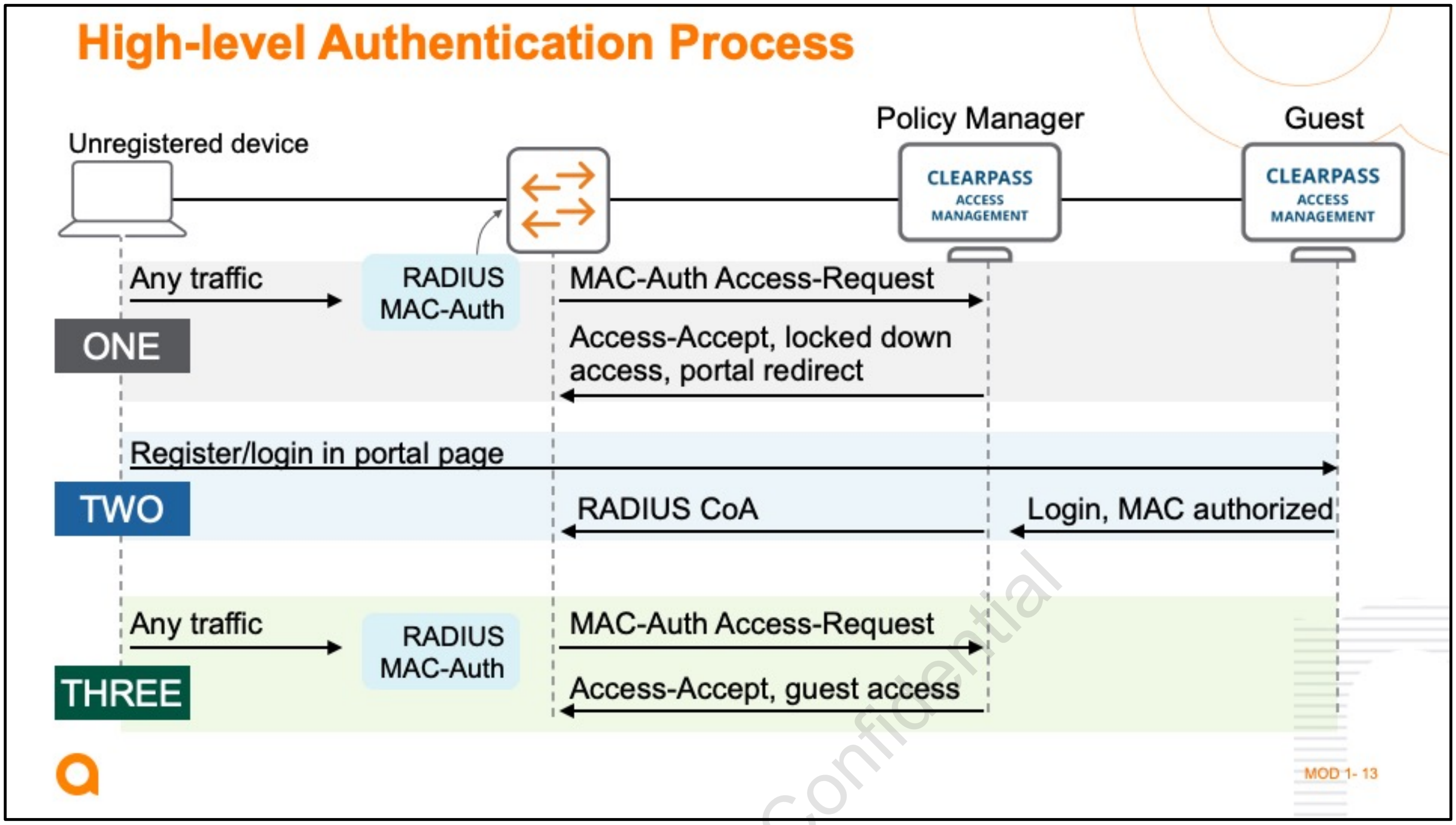


Authentication Enforced on the AOS-CX Switch

Authentication on port:
Required: RADIUS MAC-Auth
Optional 802.1X for employee onboarding

You will now look at how the captive portal solution delivers this experience. First, you must implement RADIUS-based MAC-Auth on any AOS-CX switch port to which users might connect their devices. The switch can then submit access requests for all devices to ClearPass so that ClearPass can decide how to control the devices' access.

You can optionally add 802.1X in port-based mode if you want registered devices to eventually authenticate with 802.1X. In this case, ClearPass Onboard would deploy the correct configuration settings and certificates for supporting 802.1X to devices connected to the captive portal. Companies usually use this option for employee BYOD devices, which require more security than guest devices. This module focuses more on the guest access option, though.



High-level Authentication Process

Unregistered device

Policy Manager

Guest

CLEARPASS ACCESS MANAGEMENT

CLEARPASS ACCESS MANAGEMENT

ONE

Any traffic

RADIUS MAC-Auth

MAC-Auth Access-Request

Access-Accept, locked down access, portal redirect

TWO

Register/login in portal page

RADIUS CoA

Login, MAC authorized

THREE

Any traffic

RADIUS MAC-Auth

MAC-Auth Access-Request

Access-Accept, guest access

MOD 1- 13

The figure shows the high-level connection process for an unregistered guest device, organized in three phases.

PHASE ONE

As soon as the device sends any traffic, the switch sends a MAC-Auth Access-Request with the device’s MAC address. Because the device in unknown, ClearPass sends an Accept-Accept with various dynamic settings to limit the device’s access and redirect the device to the portal page on ClearPass Guest.

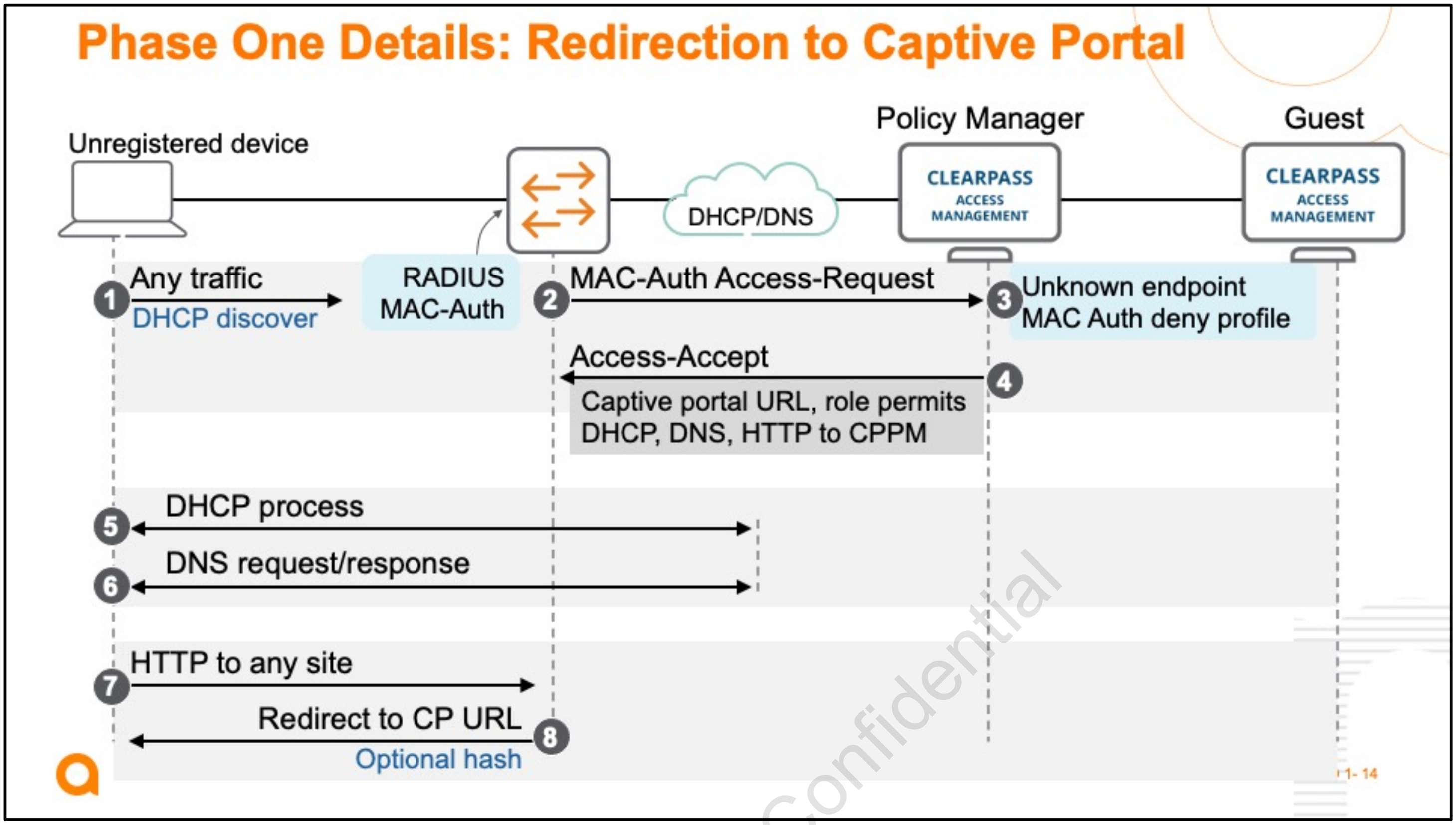
PHASE TWO

The guest uses this page to login, or to register and then login. At the end of this process, ClearPass Guest logs the user into ClearPass Policy Manager (CPPM). Policy Manager sends a Change of Authorization (CoA) to the switch.

PHASE THREE

Now when the device sends any traffic again, the switch sends a new MAC-Auth Access-Request to CPPM. Policy Manager now knows the guest device MAC address and sends an Access-Accept, with settings for appropriate guest access.

Let’s explore each of these phases in more detail.



Phase One Details: Redirection to Captive Portal

Unregistered device
Policy Manager
Guest
CLEARPASS ACCESS MANAGEMENT
CLEARPASS ACCESS MANAGEMENT
DHCP/DNS
RADIUS MAC-Auth
1 Any traffic DHCP discover
2 MAC-Auth Access-Request
3 Unknown endpoint MAC Auth deny profile
4 Access-Accept Captive portal URL, role permits DHCP, DNS, HTTP to CPPM
5 DHCP process
6 DNS request/response
7 HTTP to any site
8 Redirect to CP URL Optional hash
1-14

To improve your understanding, let's examine the captive portal redirect process in more detail. Reference the figure while reviewing the process below. In the figure, any reference to "HTTP" also includes HTTPS.

First an unregistered guest device sends any traffic, typically a DHCP discover.

The switch port enforces RADIUS MAC-Auth, so it sends an Access-Request to ClearPass Policy Manager (CPPM), which must be defined as its RADIUS server.

CPPM does not yet know the guest device MAC address. Typically, when a device cannot authenticate, the RADIUS server sends a RADIUS Reject message, and the switch port denies access. But for a captive portal solution, the ClearPass RADIUS server behaves differently, using a special "Allow All MAC-Auth" service. When a device fails authentication, this service matches the device to a "deny any" profile. The profile defines various RADIUS attributes and VSAs to apply to this device session.

ClearPass sends a RADIUS Access-Accept, with attributes to apply to the session. These attributes include:

A captive portal URL, which indicates the URL of the ClearPass portal in which users will log in or register. The URL must use standard format such as: https://clearpass.example.com. If you plan to redirect clients to an HTTPS URL, you must ensure that ClearPass has a certificate that is signed by a well-known trusted certificate authority.

A dynamic ACL with IP rules that permit only DHCP/DNS traffic, and HTTP/HTTPS traffic to ClearPass. It also permits HTTP/ HTTPS to the redirect URL.

A rate limit to reduce the risk of untrusted users or devices launching a DoS attack (this VSA is not required, but it is recommended).

- Note that you will typically use the Aruba role feature to assign these policies

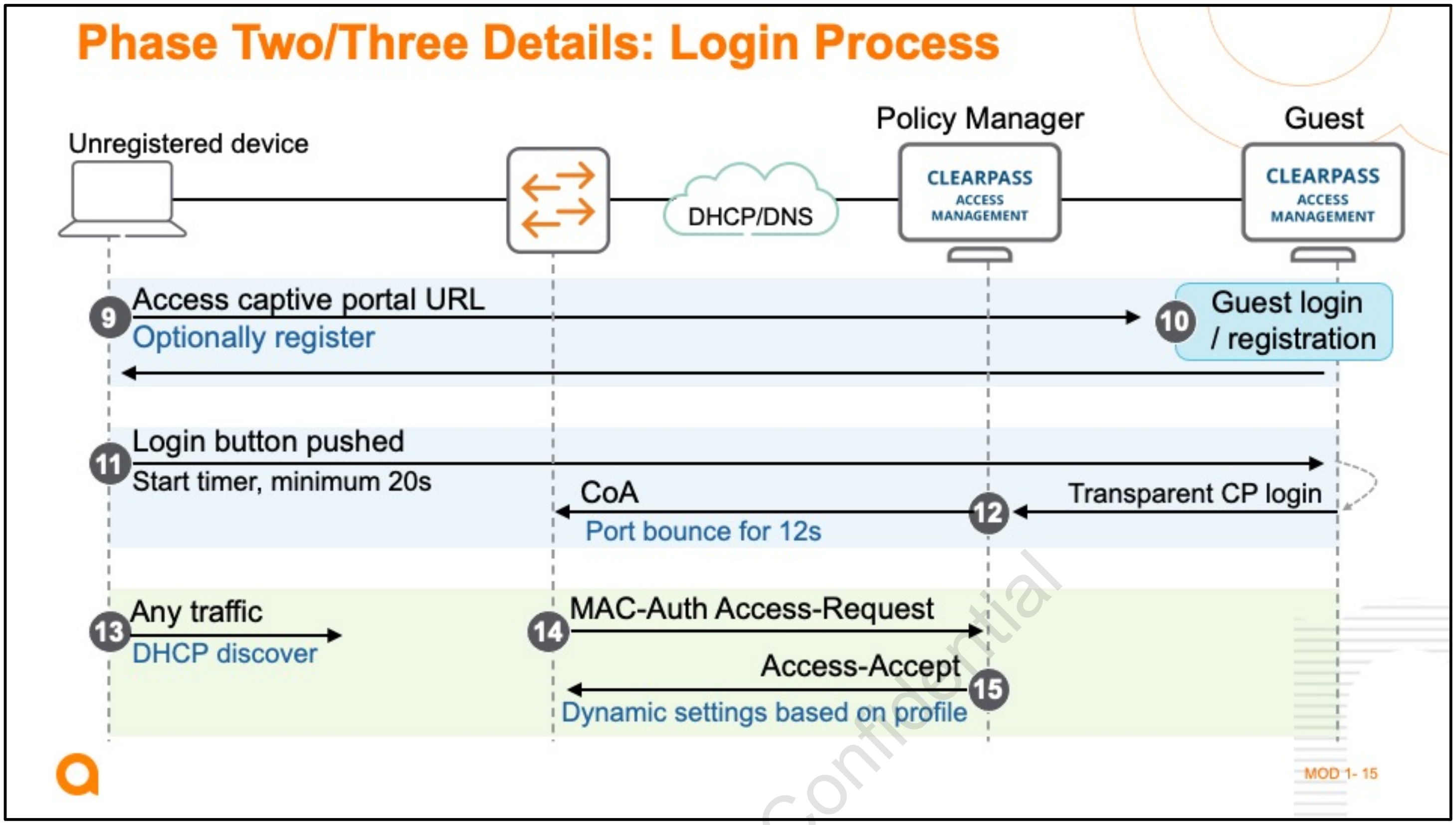
Although the device is blocked from most network access, it can send DHCP messages and receive an IP address.

And when the user opens a browser and attempts to browse to a site, the device can use DNS to resolve the hostname.

The device then sends an HTTP or HTTPS request to access some website, as desired by the user.

The AOS-CX switch intercepts that request and sends a response, posing as that address and redirecting the browser to the captive portal URL.
The switch appends some information to the URL in a query string. This information includes the device's MAC address and IP address, a timestamp, and the originally requested URL. Hackers could try to obtain inappropriate access by altering the ACL. To prevent this from occurring, the ClearPass portal can require a hash appended to any requests, created by hashing the URL with a secret key. It only responds to requests that have the correct hash and drops other requests. If this security feature is enabled on ClearPass, you must configure a key on your AOS-CX switches that matches the key on ClearPass.

The endpoint is thus redirected to the Captive Portal (CP) URL, and the process continues below.



Phase Two/Three Details: Login Process

Unregistered device

Policy Manager

Guest

DHCP/DNS

CLEARPASS ACCESS MANAGEMENT

CLEARPASS ACCESS MANAGEMENT

9 Access captive portal URL
Optionally register

10 Guest login / registration

11 Login button pushed
Start timer, minimum 20s

CoA
Port bounce for 12s

12 Transparent CP login

13 Any traffic
DHCP discover

14 MAC-Auth Access-Request

Access-Accept
Dynamic settings based on profile 15

MOD 1- 15

The unregistered device now accesses the ClearPass guest portal page. As you see, traffic flows between the client and the ClearPass Guest portal without any proxy help by the switch.

The guest might be able to register for an account based on how the portal page is set up. The ClearPass portal's appearance and functionality is highly customizable. You can configure the portal such that guests log in with an account that has been provided to them, or perhaps you want them to self-register for an account. The portal might grant all self-registering users immediate access, force them to activate the account through a text message, or force them to receive sponsor approval. ClearPass transparently adds a newly registered account to the guest repository.

When the user clicks a log in button, ClearPass Guest transparently logs the guest in to the Policy Manager. Because the wired solution uses MAC-Auth to authenticate registered devices, ClearPass Guest also makes sure that the guest device MAC address (submitted in the URL) becomes authorized. The portal page begins a login count down—recommended at about 20 seconds for a solution that uses AOS-CX switches.

How the MAC address is authorized depends on the solution. ClearPass might add the guest device to special guest endpoint repository and pair the device with the guest account. Or

ClearPass Guest might tell ClearPass Policy Manager to change the guest device, automatically discovered in ClearPass's global Endpoint Repository, to a known device. In any case, ClearPass Policy Manager now recognizes the device MAC address as a known address associated with the known guest.

The successful guest login also triggers ClearPass to send a "port bounce" Change of Authorization (CoA) message to the switch. This message causes the switch to de-authenticate the device and hold the port closed for a certain time period - 12 seconds is recommended. Note that ClearPass can be configured to use either a Disconnect Message (DM) or CoA for this purpose, but you should tell ClearPass administrators to use a port bounce CoA for AOS-CX switches.

It is important that the switch is set up to accept this message. You must enable dynamic authorization when you specify ClearPass as the RADIUS server, and you must make sure that the port is correct. The default dynamic authorization port on both AOS-CX switches and ClearPass is 3799. (If the switch does not accept the CoA message, the client will not reauthenticate, as explained in steps 13 and 14, and the client will continue to be redirected to the portal.)

When the port opens again, the device starts to send traffic (probably a DHCP Discover).

The switch port enforces MAC-Auth for the device again.

Now ClearPass finds the MAC address as a known endpoint, so it authorizes the device for access, assigning it to a profile based on policy settings. For example, it might apply a guest profile that assigns the device to a guest VLAN using the same RADIUS VSA.

Because the login counter on the portal page is several seconds longer than the port bounce time, the authentication process is complete before the countdown ends. The guest is now connected to the network and can receive the appropriate services, based on the settings applied by ClearPass.

The next time that the guest connects the device, the MAC-Auth process occurs in the background, and the guest once again receives guest-level access without having to pass through the portal. (You can set up the solution such that ClearPass forces the guest to log in through the portal again after a certain time period.)

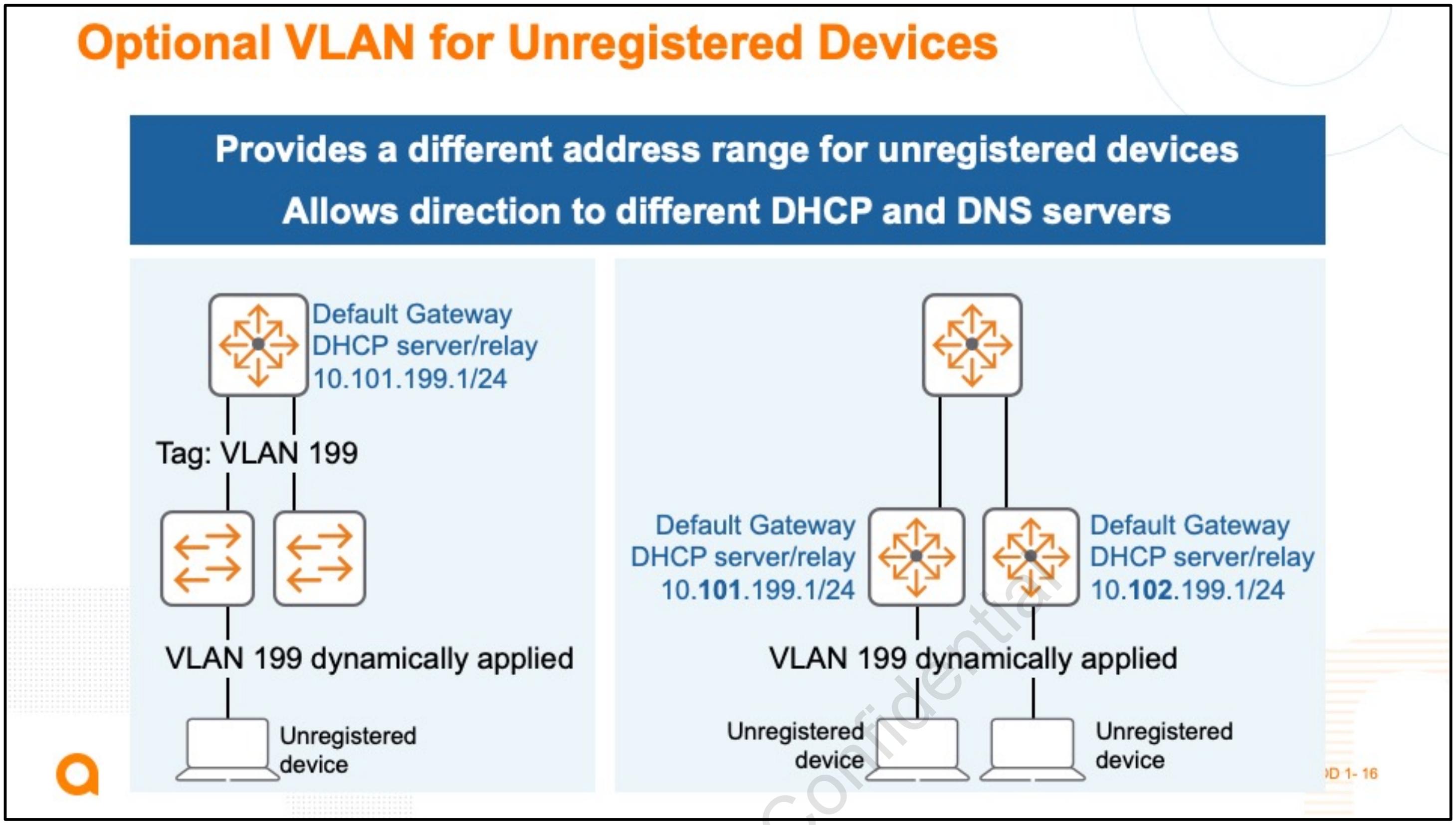
BYOD onboard

If you are using the captive portal for BYOD, you must set up the portal a bit differently. Employees typically log in to a device provisioning page on the portal with their existing domain or LDAP credentials. They then click a few buttons to accept a configuration that prepares their devices for 802.1X by configuring EAP settings and deploying certificates.

After this process, ClearPass sends a port bounce message to the AOS-CX switch port, forcing the port to close for several seconds. The device then automatically authenticates through 802.1X using the settings deployed during the provisioning process. ClearPass uses the 802.1X service to authenticate the user. The port is now open to the user with optional customization based on the ClearPass policies and profiles.

Because the switch port still enforces MAC-Auth, the switch might also submit a MAC-Auth request for the device, and the device could fail that check. However, the 802.1X authentication settings take precedence, so the device will receive the correct access and will not be redirected to the portal.

You could also choose to combine guest and BYOD solutions by adding a link to the device provisioning page on the guest login or registration page.



Optional VLAN for Unregistered Devices

Provides a different address range for unregistered devices

Allows direction to different DHCP and DNS servers

As mentioned earlier, when ClearPass sends the settings to redirect unknown devices to the captive portal, it can also send a dynamic VLAN assignment. This VLAN isolates unregistered devices in their own VLAN. Such isolation is not strictly required for security because the dynamic ACL locks down the unregistered devices' network access. However, you might prefer to place unauthenticated devices in their own VLAN for several reasons:

So unauthenticated users cannot gain information about your IP addressing scheme.

When examining traffic, you can distinguish between unauthenticated and authenticated clients.

Unauthenticated devices cannot use up IP addresses that authenticated clients require. Otherwise, hackers could attempt to obtain a lot of IP addresses to lock out legitimate users.

You can direct unauthenticated devices to different DHCP and DNS servers than authenticated users. This can help to prevent hackers from launching DoS attacks on these services.

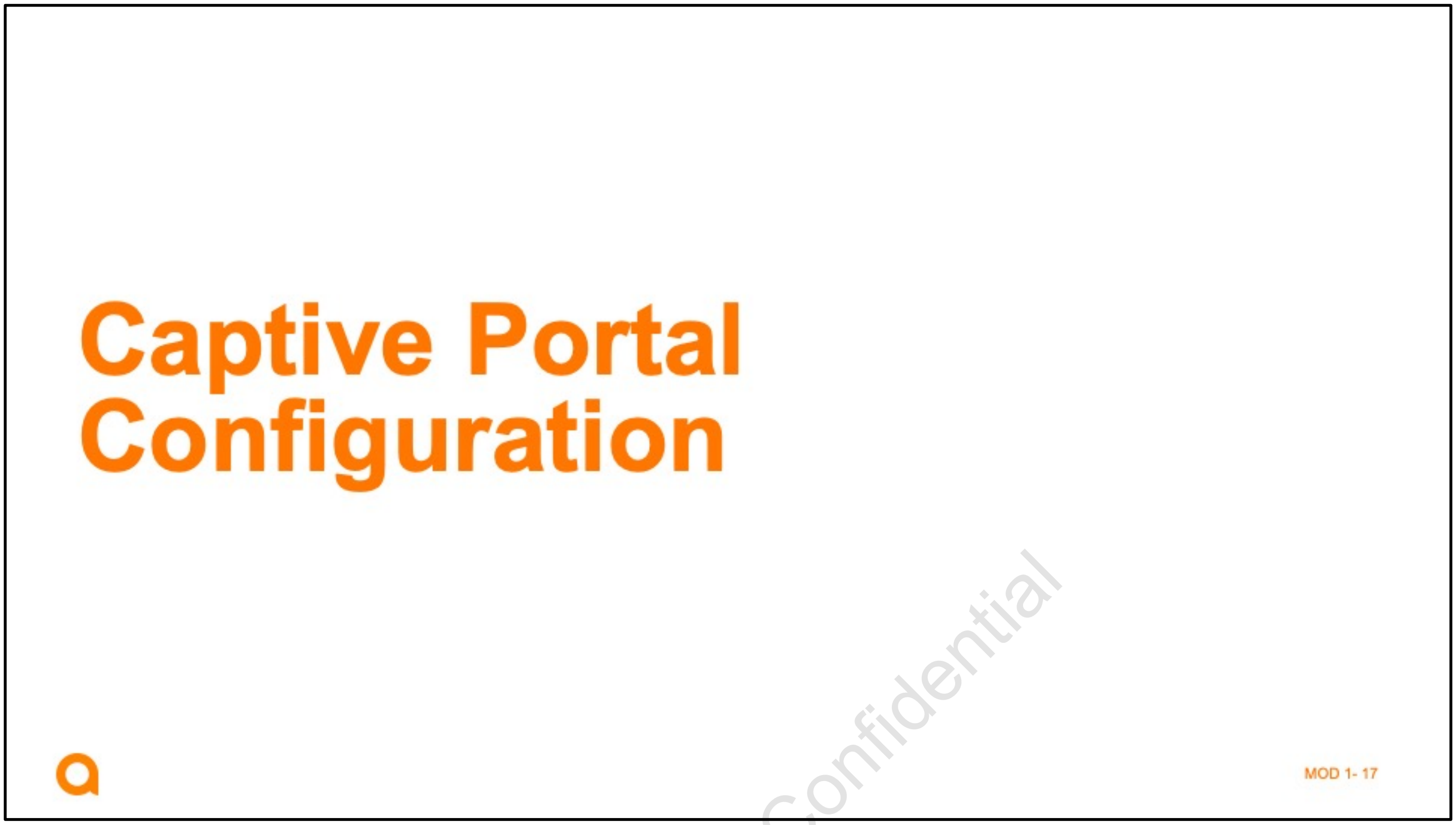
If ClearPass sends a VLAN assignment for unregistered devices, remember that you must add this VLAN to your network.

Look at the figure's left-side example. When routing occurs at the core layer, tag the VLAN on the access switch uplinks and configure the VLAN and default router IP address on the core switch (or VSF fabric).

Now look at the figure’s right-side example. When routing occurs at the access layer, each access layer switch is the default router for the unregistered VLAN. Every switch uses the same VLAN ID but must associate that ID with a different subnet. The same principle applies if you are routing at the distribution layer in a three-tier topology.

Remember to configure DHCP services on the switch or switches acting as default router. The switch can either act as the DHCP server or implement DHCP relay to a network DHCP server.

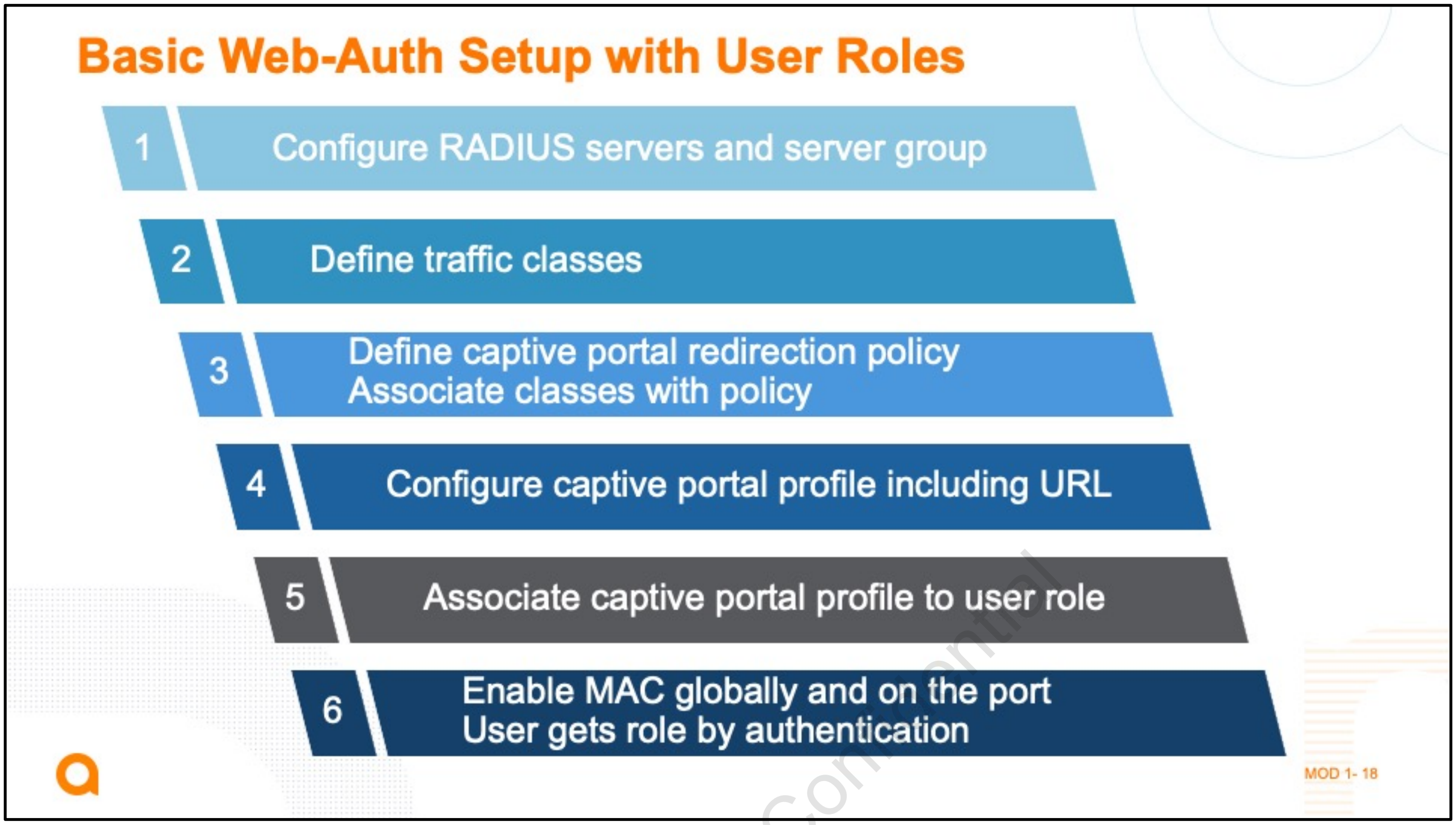
Also set up routing as appropriate to allow communication between the unregistered VLAN and the ClearPass guest server. For example, you might need to enable OSPF, preferably in passive mode, on the VLAN.



Captive Portal Configuration

MOD 1- 17

You will now learn how to implement the captive portal authentication solution, as the solution preferred by most companies seeking a full guest or BYOD access solution.



Basic Web-Auth Setup with User Roles

1 Configure RADIUS servers and server group

2 Define traffic classes

3 Define captive portal redirection policy
Associate classes with policy

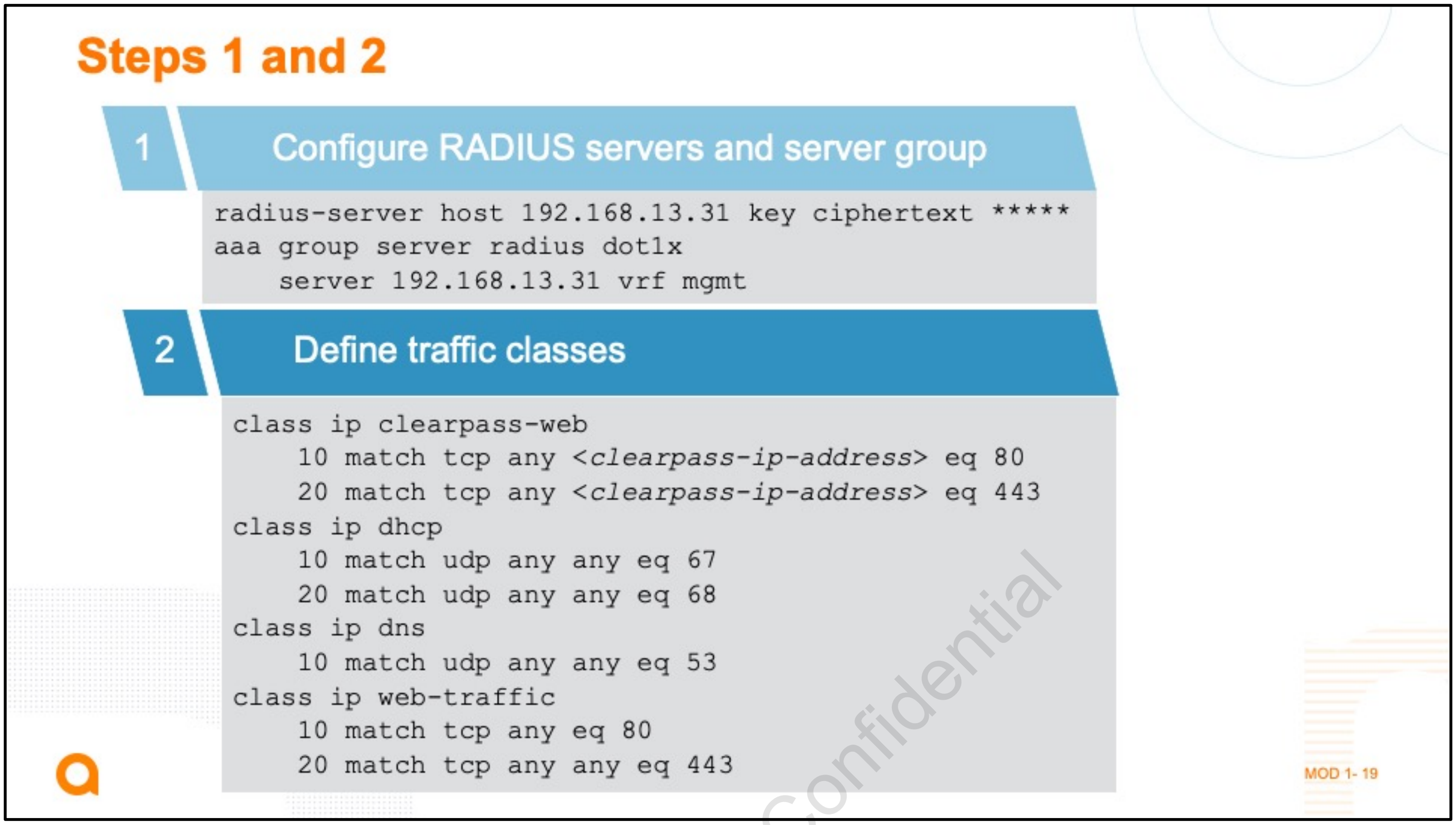
4 Configure captive portal profile including URL

5 Associate captive portal profile to user role

6 Enable MAC globally and on the port
User gets role by authentication

MOD 1- 18

Configuring an AOS-CX switch to participate in the solution is simple because ClearPass handles most of the functionality. There are six basic steps in setting up captive portal with user roles on AOS-CX switches.



Steps 1 and 2

1 Configure RADIUS servers and server group

```
radius-server host 192.168.13.31 key ciphertext *****
aaa group server radius dot1x
    server 192.168.13.31 vrf mgmt
```

2 Define traffic classes

```
class ip clearpass-web
    10 match tcp any <clearpass-ip-address> eq 80
    20 match tcp any <clearpass-ip-address> eq 443
class ip dhcp
    10 match udp any any eq 67
    20 match udp any any eq 68
class ip dns
    10 match udp any any eq 53
class ip web-traffic
    10 match tcp any eq 80
    20 match tcp any any eq 443
```

MOD 1- 19

1. Configure RADIUS servers (ClearPass) and a server group.

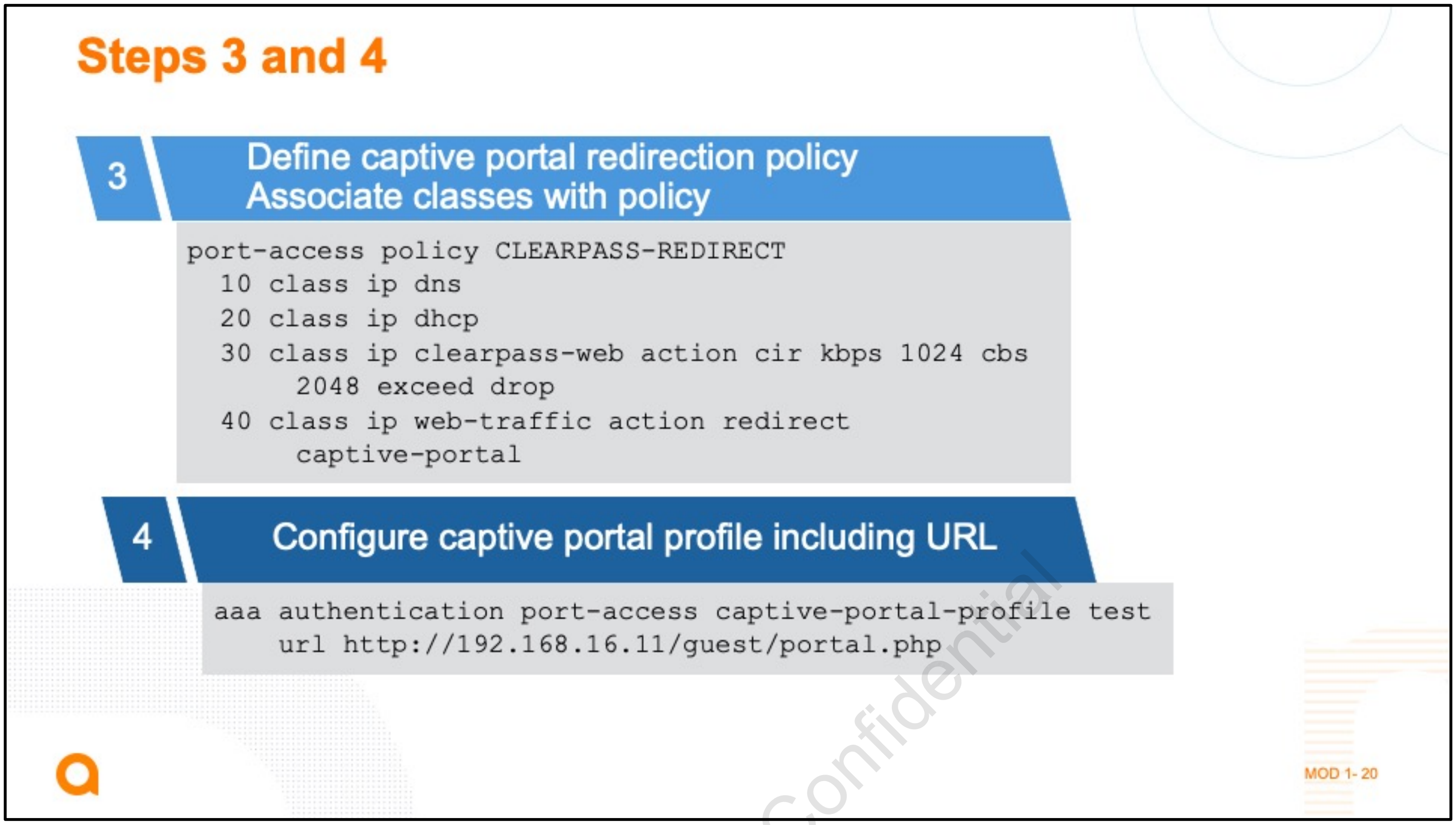
- As part of the authentication setup, make sure to specify ClearPass as the RADIUS server. Configure the correct shared secret and enable dynamic authorization. Also configure NTP as you learned how to do in previous modules.

- Also make sure that you understand whether ClearPass is deployed as a high availability cluster. The cluster provides shared configuration and databases, but not a virtual IP address. Therefore, you must make sure to specify all ClearPass servers as RADIUS servers on the switch. As you learned in a previous module, AOS-CX switches do not load balance requests between RADIUS servers. Instead, they send all requests to the first defined server unless the request to that server times out and all retries expire. If you want, you can configure one RADIUS server IP address first on half of the switches, and a different server IP address first on the other half of the switches. (You cannot define a hostname for the RADIUS server, so a DNS-based load balancer will not work.)

2. Define traffic classes.

- Traffic classes define what to match on for the policy. Traffic classes are covered in more depth in Module 15, “Quality of Service”.

- Note that the clearpass-web class map must specify the destination address of the ClearPass (or RADIUS) server that will handle web authentication.



Steps 3 and 4

3 Define captive portal redirection policy
Associate classes with policy

```
port-access policy CLEARPASS-REDIRECT
  10 class ip dns
  20 class ip dhcp
  30 class ip clearpass-web action cir kbps 1024 cbs
       2048 exceed drop
  40 class ip web-traffic action redirect
       captive-portal
```

4 Configure captive portal profile including URL

```
aaa authentication port-access captive-portal-profile test
    url http://192.168.16.11/guest/portal.php
```

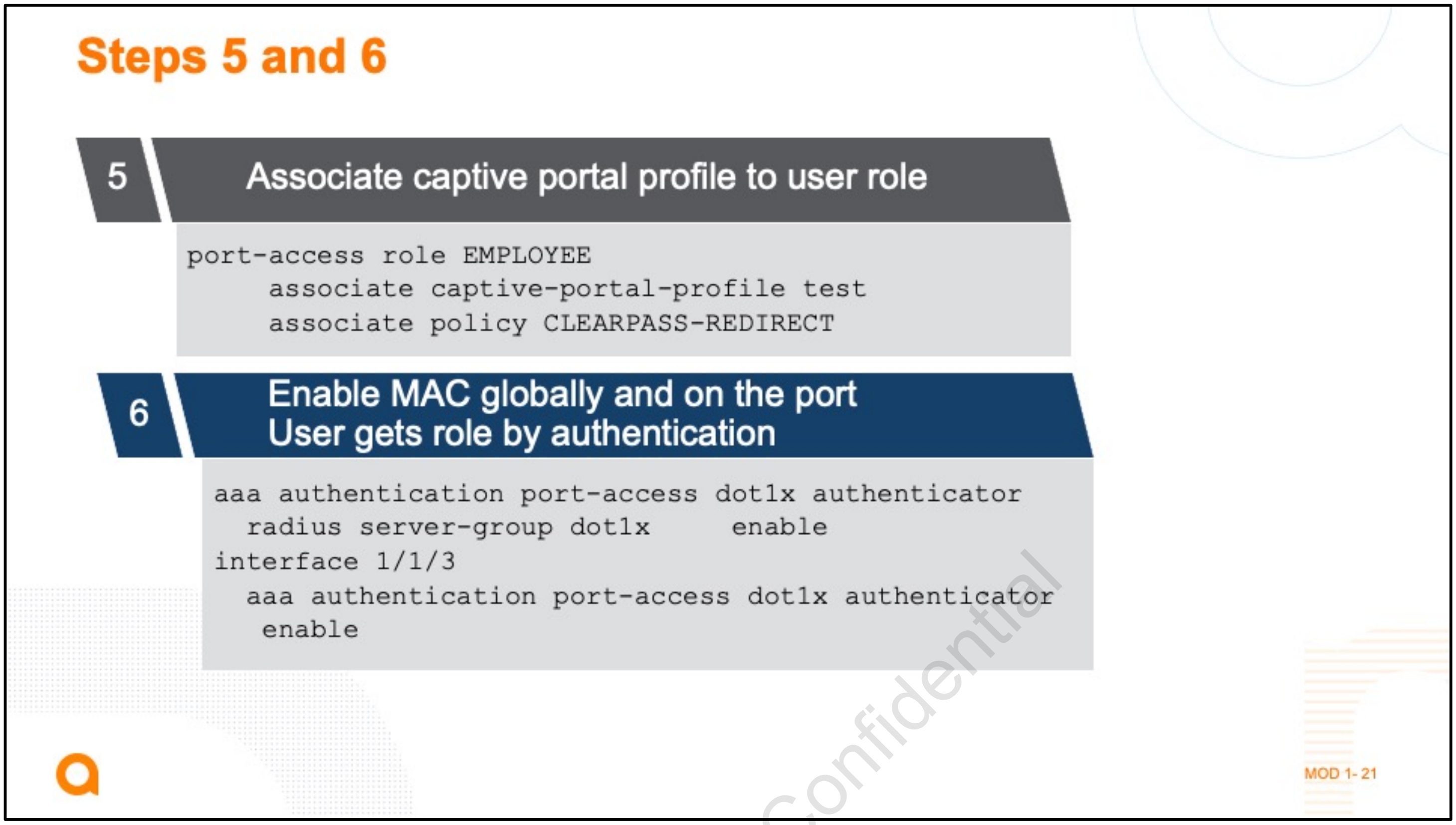
MOD 1- 20

3. Define captive portal redirection policy and associate the classes with the policy.

- Here you reference the traffic classes from Step 2. You also specify the redirect action for web traffic. Optionally you can rate limit traffic. Note that minimally you must allow DHCP and DNS, or the captive portal process will break. The redirect action in the example will redirect all web traffic.

4. Configure captive portal profile including the URL.

- In this example, the redirect URL is a ClearPass guest portal page. Optionally, you can define a hash key for protection with the url-hash-key plaintext command in the captive portal configuration context.



Steps 5 and 6

5 Associate captive portal profile to user role

```
port-access role EMPLOYEE
     associate captive-portal-profile test
     associate policy CLEARPASS-REDIRECT
```

6 Enable MAC globally and on the port
User gets role by authentication

```
aaa authentication port-access dot1x authenticator
  radius server-group dot1x     enable
interface 1/1/3
  aaa authentication port-access dot1x authenticator
   enable
```

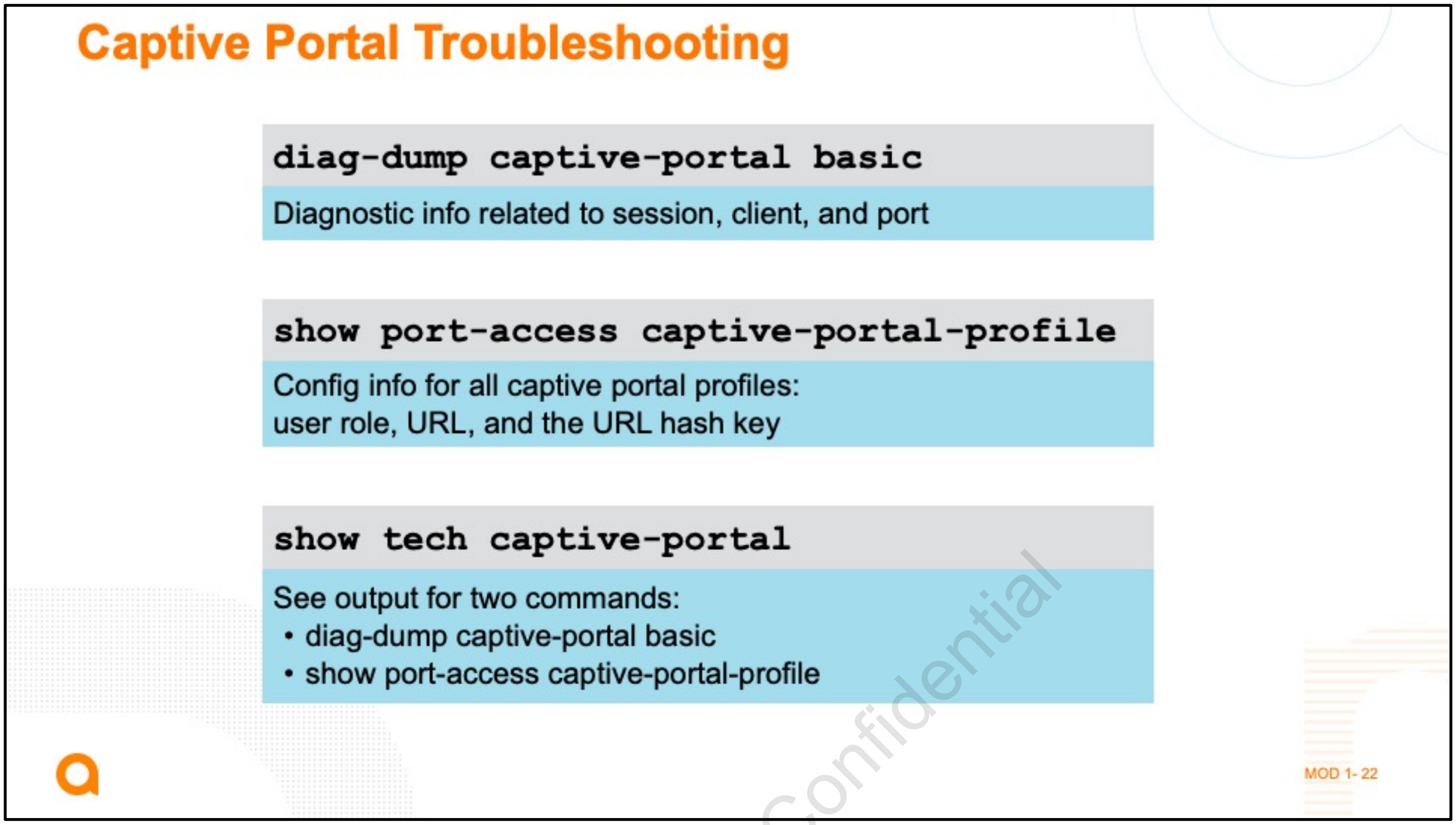
MOD 1- 21

5. Associate the captive portal profile to the user role.

- The role ties together the captive portal profile (the redirect URL) and the previously configured policy.

6. Enable MAC or 802.1X both globally and on the port to let the user get the role by authentication .

- Since captive portal is implemented with MAC authentication and RADIUS CoA on the AOS-CX switches, only MAC authentication needs to be enabled on the port. Optionally, you can enable both 802.1X and MAC-Auth if you are not sure what type of authentication will be necessary on a switch port



Captive Portal Troubleshooting

diag-dump captive-portal basic

Diagnostic info related to session, client, and port

show port-access captive-portal-profile

Config info for all captive portal profiles:
user role, URL, and the URL hash key

show tech captive-portal

See output for two commands:

- diag-dump captive-portal basic
- show port-access captive-portal-profile

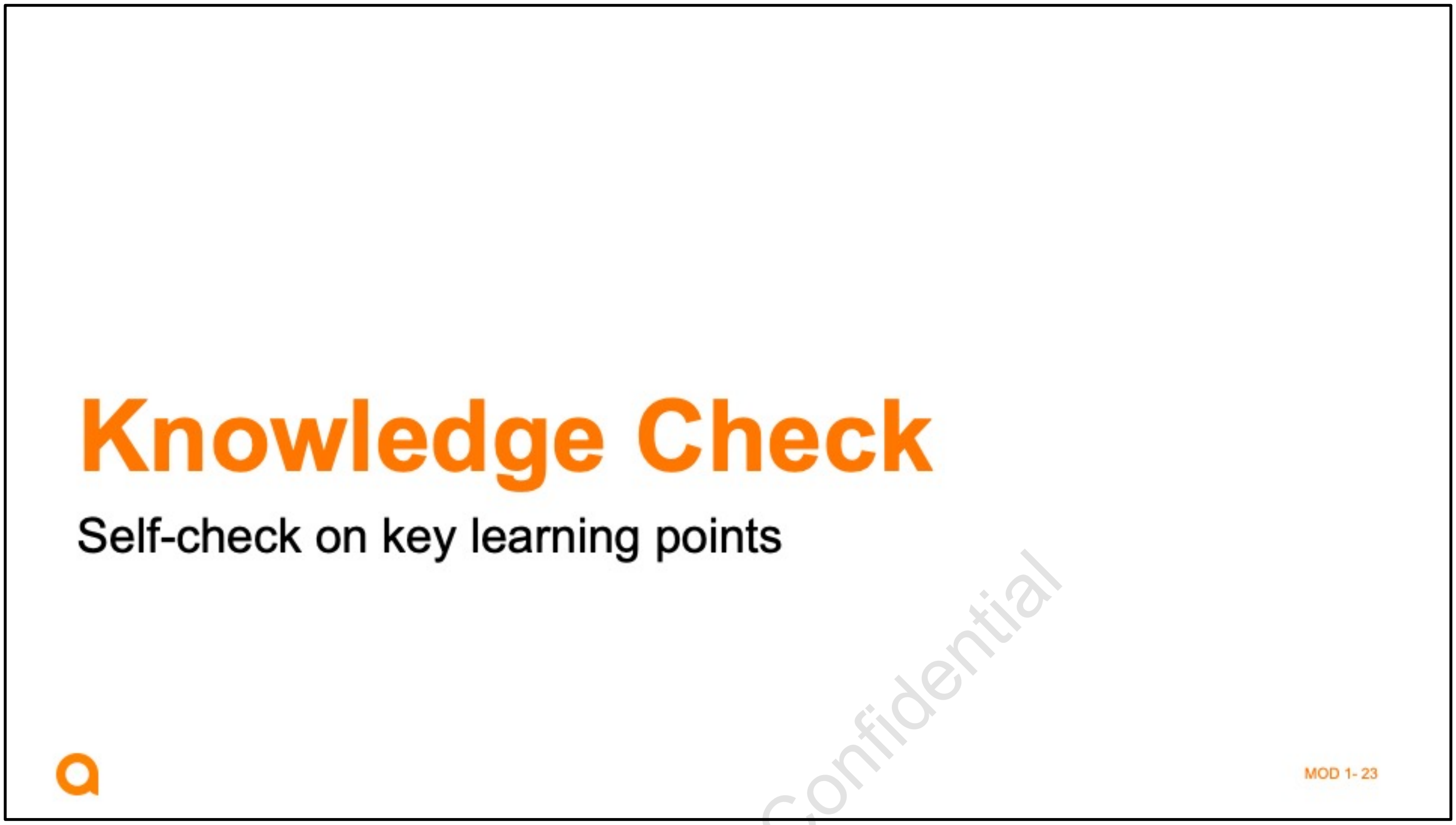
MOD 1- 22

The diag-dump captive-portal basic command displays the detailed diagnostic information related to the session, client, and port to help troubleshoot captive portal issues, including redirects.

The show port-access captive-portal profile command displays the configuration information for all captive portal profiles. Here's an example:

```
switch# show port-access captive-portal-profile
Captive Portal Profile Configuration
Name : employee
Type : local
URL : http://1.1.1.1/employee/captiveportal.php
URL Hash Key : SWNGWyMeYubHPDgVIirpEUwNK5Uf+r1vmhBIncQPw1Y=
```

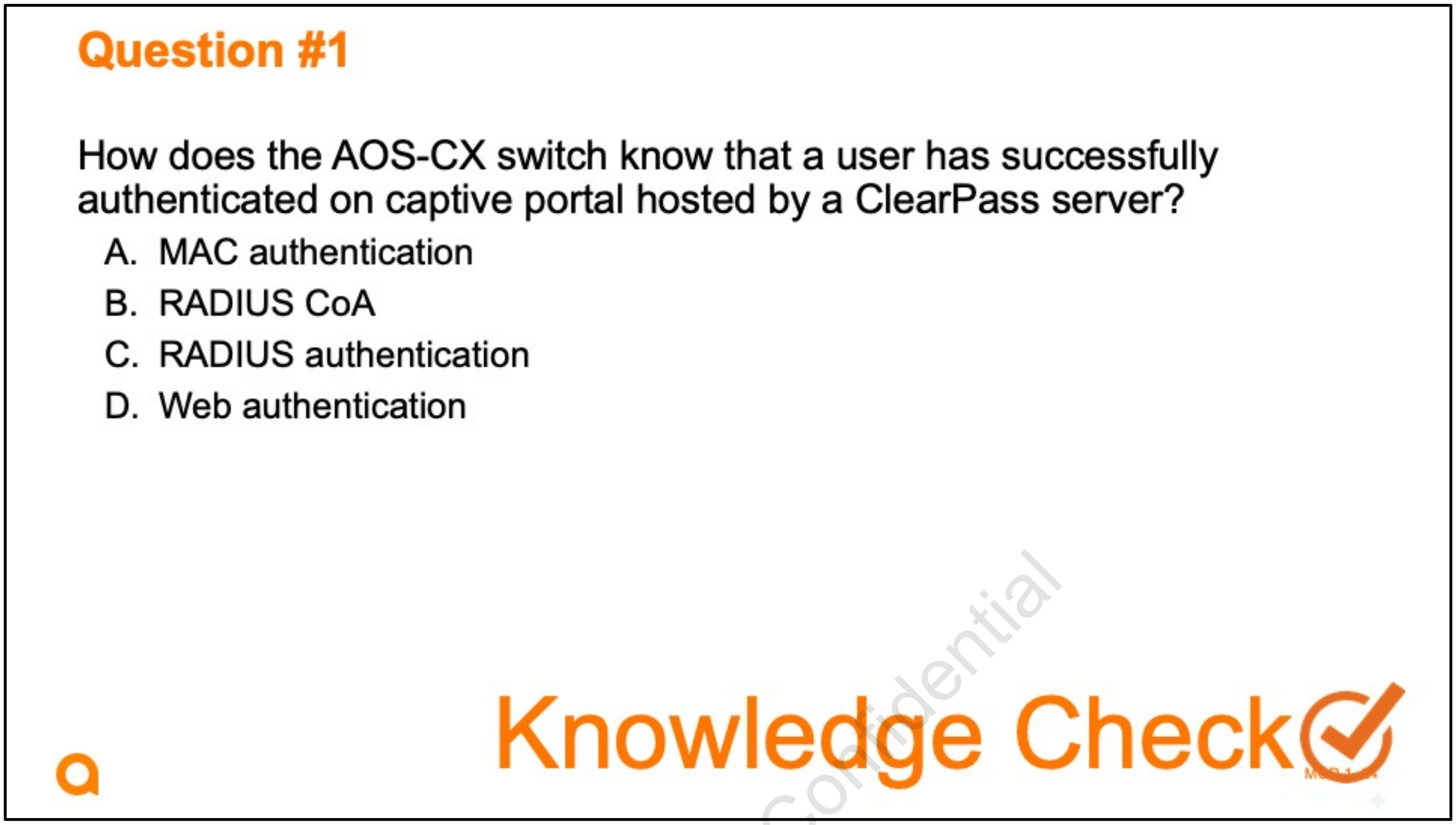
The show tech captive-portal command captures the output from the two previous commands.



Knowledge Check

Self-check on key learning points

Let’s do a knowledge check.



Question #1

How does the AOS-CX switch know that a user has successfully authenticated on captive portal hosted by a ClearPass server?

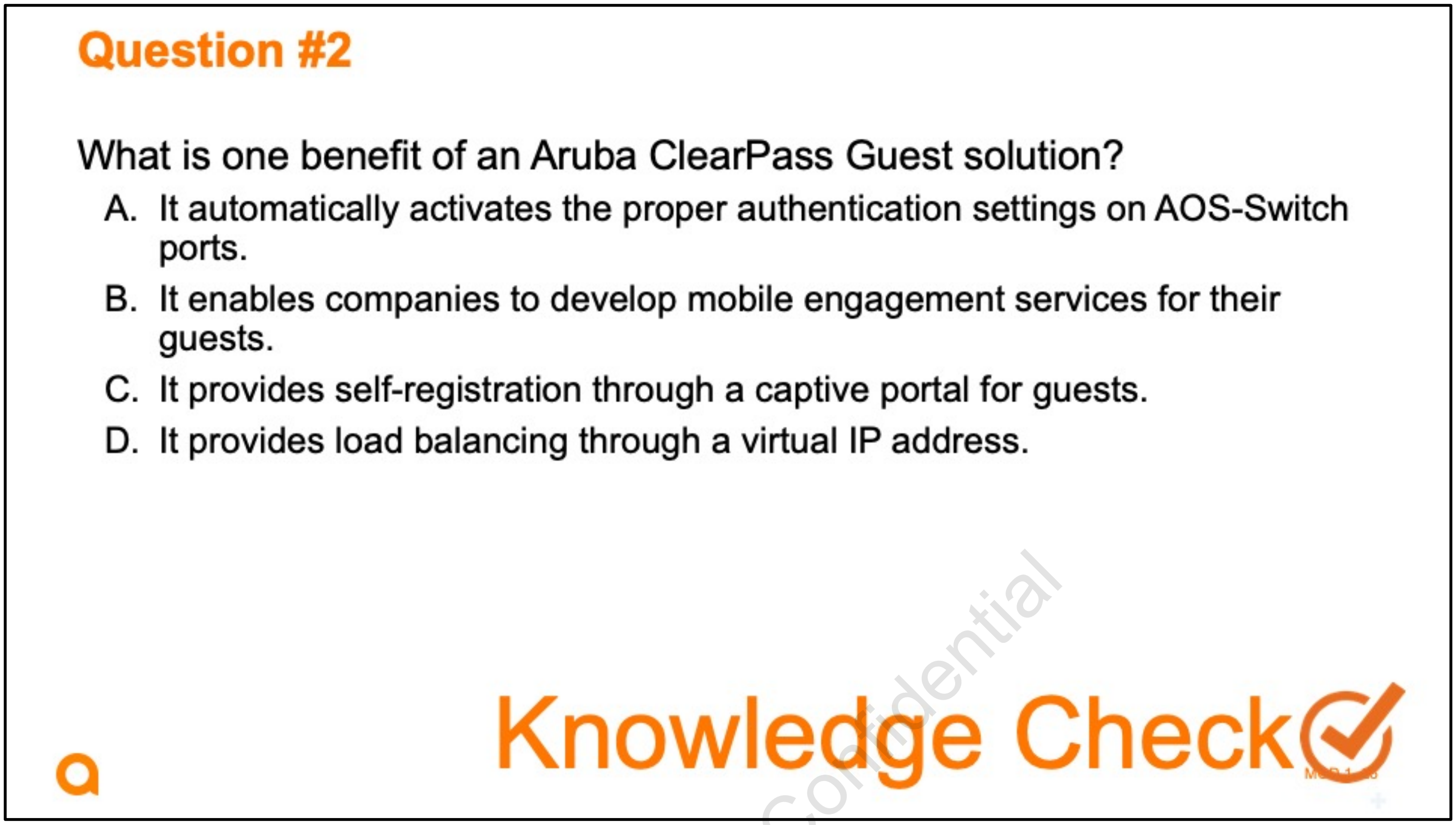
A. MAC authentication

B. RADIUS CoA

C. RADIUS authentication

D. Web authentication

Knowledge Check



Question #2

What is one benefit of an Aruba ClearPass Guest solution?

A. It automatically activates the proper authentication settings on AOS-Switch ports.

B. It enables companies to develop mobile engagement services for their guests.

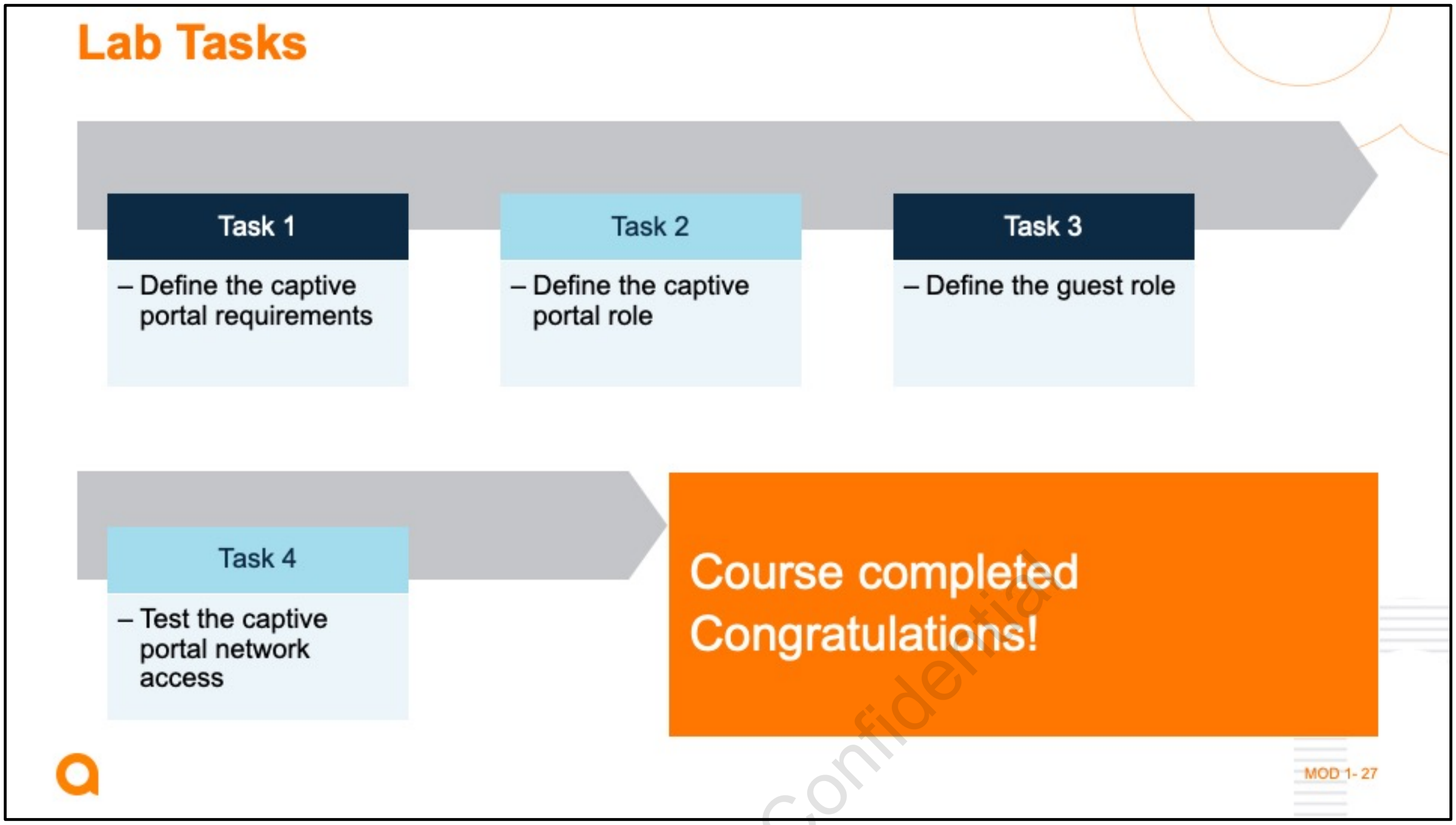
C. It provides self-registration through a captive portal for guests.

D. It provides load balancing through a virtual IP address.

Knowledge Check

Lab Activity

One more lab activity.



Lab Tasks

Task 1

– Define the captive portal requirements

Task 2

– Define the captive portal role

Task 3

– Define the guest role

Task 4

– Test the captive portal network access

Course completed
Congratulations!

The figure provides a brief review of lab tasks. Please see your lab guide for details. Once you have completed the lab and knowledge check, you have completed this course. Congratulations!

3333 Scott Blvd, Santa Clara, CA 95054
TEL: 408.227.4500 | FAX: 408.227.4550
www.ARUBANETWORKS.com

EDU-ICX-ILT-v20.211